MANAGING SPACE IN FORWARD PICK AREAS OF WAREHOUSES FOR SMALL PARTS

A Thesis
Presented to
The Academic Faculty

by

Sriram Subramanian

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Industrial and Systems Engineering

Georgia Institute of Technology
December 2013

MANAGING SPACE IN FORWARD PICK AREAS OF WAREHOUSES FOR SMALL PARTS

Approved by:

Professor J.J. Bartholdi III
Advisor
School of Industrial and Systems Engineering
Georgia Institute of Technology

Professor Steven T. Hackman
Advisor
School of Industrial and Systems Engineering
Georgia Institute of Technology

Professor John H. Vande Vate
School of Industrial and Systems Engineering
Georgia Institute of Technology

Professor Alan L. Erera
School of Industrial and Systems Engineering
Georgia Institute of Technology

Professor Kevin R. Gue
Department of Industrial and Systems Engineering
Auburn University

Date Approved: 22 October 2013

To

the memory of my grandfather,

Dr. A.S. Ramanathan.

ACKNOWLEDGEMENTS

Writing this thesis is one of the toughest things I have done. And it was made possible only because of the support and guidance of my advisors, teachers, friends and family. I want to thank Professor John J. Bartholdi III for his guidance and wisdom. He exposed me to new research areas, brought clarity to my thinking and more importantly, supported and guided me in every aspect of the process. I also have to thank Professor Steven T. Hackman who helped transform what was a rough draft into a thesis by guiding me on formulating and writing the ideas in a coherent manner. I am grateful to Pete Viehweg for spending time with me to discuss my research ideas on warehouses and his insightful comments that helped refine them in a significant manner. I also have to thank Professor Jim Wilson of North Carolina State University who exposed me to research and encouraged me to pursue a doctoral program. Many thanks to my committee members, Professors J.H. Vande Vate, A.L. Erera and K.R Gue for serving on my committee and offering valuable comments. I am also grateful to Professor Sridhar Seshadri of Indian School of Business who has helped me focus on writing this thesis during other projects we have worked on together. I am greatly indebted to many Professors who offered interesting classes at Georgia Tech that have shaped my thinking. Many concepts and ideas that I have applied in this thesis and remember today are due to their ability to teach in a clear and concise manner.

Over the years I have had great support from many of my fellow students who have worked with me on class projects, discussing and understanding ideas, concepts and techniques, and who over time, have become my good friends. Thank you (in alphabetical order) to Gauthaman Balaraman, Yun Fong Lim, Marcos Goycoolea, Yongpei Guan, Vivek Gupta, Yetkin Ileri, Jim Luedtke, Anuj Shah, Ramesh Singh and Krishna Tunga. There are many others whose names I have missed out who have helped me along the way and I am very grateful to them.

I have to thank my family who have been with me every step of the way. They have supported me through the ups and downs of graduate school. Most importantly, I have to thank my loving and patient wife, Divya, who in addition to supporting and encouraging me, helped me with may have seemed to be an almost endless rounds of proof-reading this thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Many high-volume warehouses for small parts such as pharmaceuticals, cosmetics and office supplies seek to improve efficiency by creating *forward pick areas* in which many popular products are stored in a small area that is replenished from reserve storage. This thesis addresses the question of how to stock forward pick areas to maximum benefit. Achieving this maximum benefit requires answers to two key, inter-related decisions:

Assignment Which SKUs should be stored in the forward pick area?

Allocation How much space should be allocated to each SKU?

With respect to the allocation problem, we introduce a *Powers of Two* allocation scheme designed to simplify shelf management. It restricts allocations to values that are a small constant times a power of two. The scheme is very effective and we prove an *a priori* bound of 6% above the optimal value. The scheme results in allocations that are manageable. It generates very few unique values of allocations and the allocations share space and restock activity much more evenly than conventional stocking strategies. For warehouses with many SKUs, we furthermore prove that *almost* any constant can be chosen and the solution quality is still guaranteed to be within 12.5% of optimal value. To help warehouse managers synchronize restocking activity, we show how to allocate space so that the frequencies with which SKUs must be restocked are restricted to powers of two.

We present a ranking-based algorithm to assign and allocate SKUs to warehouses with multiple forward pick areas. It produces a near-optimal solution, it is easy to implement, and it is computationally efficient. We show that a similar algorithm

generates a near-optimal solution for extensions that account for constraints on congestion within the forward pick area, as well as constraints on number of picks or restocks. We also show how to determine the optimal assignment for warehouses with one or more forward pick areas that allocate space in two ways that are common in practice: the allocation of identical amounts of space to each SKU (Equal Space) and the allocation of space to store equal time supplies of each SKU (Equal Time). We also consider extensions to account for constraints on total number of picks and restocks, and discuss solution methodologies.

We critically examine one of the rules-of-thumb by which warehouses have traditionally been managed: the *80-20* rule, also known as *ABC* distribution, or more formally, the *Power Law*. Power laws have been identified in many natural processes, from populations of cities to number of visits to websites. Warehouses frequently use the 80-20 rule to characterize the popularity of SKUs as reflected in the number of requests (picks). We examine the empirical picks-per-SKU distributions from thirty warehouses stocking different classes of items, from apparel to service parts for trucks. We fit a power law distribution to the data, check if the fitted distribution is indeed suitable, and test other candidate distributions. We find that out of the thirty data sets, sixteen exhibit a power law distribution in the upper tail of picks. However, out of these sixteen data sets, twelve show significant statistical support for alternative distributions and only in four did the power law turn out to the best fit. We test the hypothesis that the picks-per-SKU distributions of warehouses in similar industries are themselves similar. We find that the fitted distributions vary significantly within industry categories. We review explanations for why power laws arise in other settings, and identify explanations that are plausible in the warehouse setting.

CHAPTER I

INTRODUCTION TO FORWARD PICK AREAS

1.1 Forward pick areas in warehouses for small parts

Order picking is the single most expensive activity in a warehouse, accounting for approximately half of the operating costs [13]. As warehouses diversify their product lines and at the same time allow customers to order small quantities, the efficiency of order picking has become vital in warehouse operations. Efficiency in picking is measured in two ways: the labor costs of picking, and responsiveness to customer demand by quick fulfillment of orders. To that end, many high-volume warehouses seek to improve the efficiency of picking by creating *forward pick areas*. Forward pick areas concentrate picks in a small area, thereby reducing travel for order pickers. This enables quick and low cost picking as well as closer supervision of order pickers. Forward pick areas are particularly well suited to warehouses for small parts such as pharmaceuticals, cosmetics or office supplies. These warehouses typically stock a large assortment of items that are also small enough to store in sufficient quantity in a small area so as to concentrate picks.

A warehouse may have a forward pick area and sometimes even multiple forward pick areas apart from the *reserve*, or main storage area. The reduced travel in the forward pick area facilitates picking at a lower cost relative to the reserve storage. Many of the popular Stock-Keeping Units (SKUs) are stored in the forward pick area in small quantities, so that picking can be accomplished in a small area. This is typically achieved by employing specialized storage mechanisms such as flow racks or carousels that allow high storage densities, making it easier to pick by reducing travel time for order pickers between SKU locations. Figure 1 shows the pick face of a gravity flow rack, a type of equipment that is commonly deployed to create forward pick areas. The flow rack is picked from the front and restocked from the back so as to keep interference between those activities to a minimum. The inclined shelf and rollers allow the force of gravity to automatically move the cases to the front, allowing the picker to easily reach for the carton. Furthermore, if the warehouse

Figure 1: The pick face of a gravity flow rack. The shelf is inclined with rollers so that the cartons automatically move to the front due to gravity. The flow rack is restocked at the back of the shelf so as to not to interfere with picking. Courtesy of `www.warehouse-science.com`.

layout allows fast travel to shipping from the forward pick area, the order fulfillment process becomes that much quicker. This means, however, that the forward pick area has limited space and cannot store the entire inventory. Therefore, the SKUs have to be replenished from the much larger reserve, incurring restocking costs. The decision on how to stock the forward pick area hence involves balancing a lower cost per pick against restocking costs required to sustain the forward pick area, as illustrated in Figure 2. Most of this thesis focuses on addressing the question of how to stock the forward pick area to maximum benefit by using mathematical models.

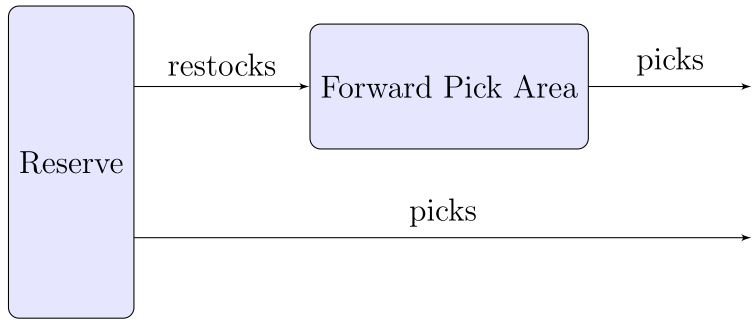

Figure 2: A warehouse with one forward pick area. The forward pick area enables low cost picks relative to the reserve, but needs to be restocked periodically from the reserve.

Achieving this maximum benefit requires two key, inter-related decisions:

Assignment Which SKUs should be stored in the forward pick area?

Allocation How much space should be allocated to each SKU?

Re-configuring the SKUs in the forward pick area by changing the assignment or allocation incurs a significant cost. Therefore, warehouses have to decide on assignment and allocation and maintain the configuration for a given time period, the planning horizon. They base their configuration on expected orders (number and quantity) for SKUs, and any variation in realized orders causes inefficiency in the utilization of the forward pick area.

The space allocated to each SKU is constrained because of the geometry of the SKUs and the storage units in the forward pick area. For example, the number of cartons that can be stored depends on the available height and depth in a shelf, along with carton dimensions. Warehouse layouts, in addition to constraining space available in the forward pick area, may also place limits on the rate of material flow between the reserve and the forward pick area, and between the forward pick area and shipping. The labor force and equipment available also imposes restrictions on picking and restocking activities in a given time period. For example, given the number of order picking staff, the warehouse can only pick a certain number of orders from the forward pick area in a day, placing a limitation on the assignment of SKUs to the forward pick area. Such restrictions on restocking impose an indirect constraint on how much of a SKU can be stored in the forward pick area, as that determines the number of restocks.

Within the limitations and constraints imposed by the warehouse, managers strive to make the trade-off between pick savings produced by storing SKUs in the forward pick area, and the resulting restocking costs. In practice, however, most warehouses use rules-of-thumb to make decisions on fast pick areas and this often results in inefficient usage of the forward pick areas. For example, typically, they assign their most popular SKUs, i.e. the SKUs expected to generate the most number of picks, to the forward pick area. To allocate space, they almost always employ one of two following stocking strategies, as found by Bartholdi and Hackman [8] and Van den Berg et al. [39] in their surveys.

1. *Equal Space* strategy allocates equal space to all SKUs assigned to be stored in the forward pick area.

2. *Equal Time* strategy stores enough supply of each SKU such it is expected to fulfill the demand for an equal period of time. This implies that all SKUs are restocked at the same time and therefore that each SKU is expected to incur equal number of restocks during the planning horizon.

Both strategies result in sub-optimal utilization of the forward pick area [8].

Mathematical models designed to construct the optimal allocations of space in a forward pick area must consider the following factors. The model must consider the SKUs to stock and their respective pick and restock costs (which mostly depend on travel time), and the storage capacity in each of the available forward pick areas. The model also has to consider the stocking strategy (e.g. Equal Space) that is acceptable for the warehouse. Finally, to optimize expected future costs of picking from and restocking the forward pick area, the model has to account for the forecast of number of orders and quantities by SKU for the planning horizon.

Hackman and Rosenblatt [23] were the first to describe a mathematical model for the optimal assignment and allocation of SKUs that minimizes the combined costs of picking and restocking in a warehouse for small parts with one forward pick area. They propose a fluid model that treats the volume occupied by each SKU as continuously divisible and incompressible. In the next section, we review the Hackman and Rosenblatt [23] fluid model and related developments in the literature.

1.2 The fluid model for space allocation

The fluid model treats each SKU as an incompressible, continuously divisible fluid instead of the discrete units of volume that it actually is, such as a pallet, case or each. This is a reasonable assumption in the case of a warehouse for small parts where the individual SKUs are relatively small in size.

Using this model, we define *flow* as the volume of product forecast to move through the warehouse in a time period. For a SKU i, with flow f_i per year, and volume of space allocated in the forward pick area v_i, the number of restocks forecast will be

$$\frac{f_i}{v_i} \text{ restocks per year.}$$

We make the following assumptions on the restocking model, as stated in Bartholdi and Hackman [8].

A1 When the pick quantity exceeds the amount available in the forward pick area, a restock is triggered.

A2 The pick quantity never exceeds the full allocation of a SKU in the forward pick area. In practice, unusually large quantities are typically filled from the reserve.

A3 The entire quantity to be restocked can be brought in one trip.

The restock cost is based primarily on the time to travel from the forward pick area to the reserve and back. In addition, we must consider the following four cost components of restocking, as discussed in Bartholdi and Hackman [8].

- *The time needed to travel within the forward pick area to the SKU location.* This travel cost is small as the entire space occupied by the forward area is small in comparison to the reserve.

- *The time needed to travel between the forward pick area and the reserve.* This travel cost depends on the warehouse layout and not on the locations of the SKUs to be restocked.

- *The time traveled within the reserve storage to the location of the SKU to be restocked.* This travel cost is variable, as random storage is typically employed for the reserve, and an average travel time can be used to represent this time regardless of the SKU.

- *The cost of handling the units when restocking.* This handling cost is incurred for every unit shipped and hence is fixed with respect to the quantity to store in the forward pick area.

In each case, the additional cost components are either fixed or small. Therefore, the total restock cost is proportional to the total number of restocks made, which is independent of the SKU.

Under these assumptions, Hackman and Rosenblatt [23] pose the problem of allocating space among a given set of SKUs, $N = \{1, 2, \ldots, n\}$, so as to minimize the total number of restocks, as shown in (1.2.1). The decision variable v_i is the space allocated to SKU i in the forward pick area.

$$\min Z(\mathbf{v}) \equiv \sum_{i \in N} f_i / v_i \tag{1.2.1a}$$

$$\sum_{i \in N} v_i \leq V_F \tag{1.2.1b}$$

$$v_i > 0, \qquad\qquad i \in N$$

The constraint (1.2.1b) ensures that the sum of the space allocated to the SKUs in forward pick area F is less than the capacity, V_F. In fact, the nature of the number of restocks defined with the volume allocated to each SKU in the denominator ensures that the capacity constraint is tight at optimality. By applying standard Lagrangian techniques for solving constrained convex minimization problems, Hackman and Rosenblatt [23] determine the optimal solution $\tilde{\mathbf{v}}$ to be:

$$\tilde{v}_i = \frac{\sqrt{f_i}}{\sum_{j \in N} \sqrt{f_j}} V_F. \tag{1.2.2}$$

Bartholdi and Hackman [8] justify the use of the Hackman and Rosenblatt [23] fluid model shown in (1.2.1) to allocate space in a warehouse for small parts. Applying the same fluid model approach, they determine the corresponding allocations for Equal Space and Equal Time strategies as follows. In Equal Space, each SKU is allocated the same amount of space and so,

$$v_i = \frac{V_F}{|N|}.$$

Similarly, the Equal Time allocation is such that each SKU has the same number of restocks, which implies that

$$v_i = \frac{f_i}{\sum_{j \in N} f_j} V_F.$$

They show that the Equal Time and Equal Space are identical in the total number of restocks required. Moreover, the two strategies cost significantly more than the optimal allocation, in terms of number of restocks required to sustain the forward pick area for typical warehouses. The optimal allocation is also more *manageable* as it shares space among SKUs more evenly than Equal Time, and shares restocks more evenly than Equal Space. Bartholdi and Hackman [8] extend the fluid model for optimal allocations to allow the restock cost to vary by SKU, and to account for reorder points and safety stock.

The simplicity of the fluid model, however, does not permit consideration of the geometries of SKUs and storage units in the forward pick area. Recently, Walter et al. [40] using the fluid model in (1.2.1) show how to allocate space in a forward pick area with equal size storage units, where each SKU can be allocated only a discrete number of storage units. For a given assignment of SKUs, they show how to determine the optimal allocation such that every SKU has an allocation $v_i \in \{1, 2, \ldots, V_F\}$, where the capacity V_F is expressed in number of the storage units. This type of allocation is practical in the case of a flow rack with equal sized lanes where each SKU can only be allocated a discrete number of lanes.

1.3 Assignment and allocation of SKUs

The previous section shows the representation of the fluid model to allocate space in a warehouse with one forward pick area and a given assignment of SKUs. In this section, we review the full Hackman and Rosenblatt [23] model that includes the problem of assignment and allocation of SKUs.

The assignment of a SKU to the forward pick area lowers pick costs by a fixed amount for each pick made, but causes restocking costs. From the previous section, the restocking cost is assumed to be fixed per restock and is denoted by c_r. We assume that the lower cost per pick is mainly due to the difference in travel time to the forward pick area and the reserve, on the basis of the following analysis of the components of the pick cost. The pick cost includes the cost of travel to the SKU location and back, plus the cost of handling the units ordered. The handling cost depends on the total demand and is independent of the quantity or the location of the SKU. The primary cost is thus the cost of travel to the SKU location. Making the same assumptions on the piece picking travel cost within the forward pick area and the reserve as those made for restock costs, we see that the primary measure of pick cost is the travel to the forward pick area or the reserve, depending on where the SKU is to be picked from. The difference between the cost of travel to the reserve and the cost of travel to the forward pick area is the pick cost savings. Let this difference in cost of travel per trip be s. If the forecast number of picks for SKU i is p_i, then the net savings on picking cost for a SKU is sp_i. Figure 3 illustrates the cost model of a SKU assignment to the forward pick area.

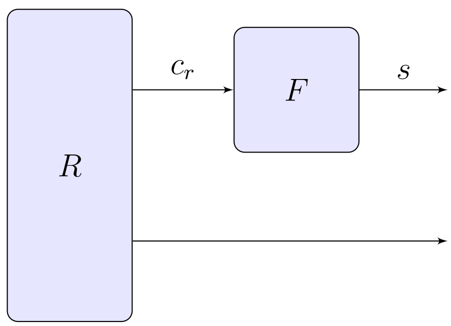

Figure 3: A warehouse with one forward pick area, F that has a pick savings of s compared to the reserve, R. The cost of restocking the forward pick area from the reserve is c_r.

Following Hackman and Rosenblatt [23], we define the decision variables as follows. Let v_i be the volume of space in F allocated to SKU i and x_i be a binary variable that is 1 if SKU i is assigned to F and 0 otherwise, for all $i \in N$. The objective is to maximize the total *net benefit*, which is the savings in picking costs minus the restocking costs as shown in formulation (1.3.1).

$$\max_{\mathbf{x},\mathbf{v}} \sum_{i \in N} x_i \left(sp_i - \frac{c_r f_i}{v_i} \right) \tag{1.3.1a}$$

subject to

$$\sum_{i \in N} x_i v_i \leq V_F \tag{1.3.1b}$$

$$x_i \in \{0, 1\} \qquad i \in N \tag{1.3.1c}$$

$$v_i \geq 0 \qquad i \in N \tag{1.3.1d}$$

To solve the formulation (1.3.1), Hackman and Rosenblatt [23] propose a heuristic procedure to determine the SKU assignment on the basis of ranking SKUs by the ratio of expected number of picks to square root of the volume of product forecast to be shipped in the demand period. That is, a sku i is said to outrank sku j if:

$$\frac{p_i}{\sqrt{f_i}} \geq \frac{p_j}{\sqrt{f_j}}$$

They rank the set of SKUs in descending order of the ratio, assign the k highest ranked SKUs to the forward pick area and evaluate the total net benefit for $k = 1, 2, \ldots, |N|$. The value of k that corresponds to the solution with the highest total net benefit is chosen. This contrasts with the conventional assignment method of choosing SKUs for the forward pick area based on popularity. In a set of numerical examples, they show that the heuristic outperforms the conventional method by a significant margin. In addition, they derive a set of *a priori* and *a posteriori* conditions on the data, from which they prove that the heuristic determines the optimal solution if either set of conditions hold.

Frazelle et al. [14] generalize the Hackman and Rosenblatt [23] model for the warehouse with one forward pick area to simultaneously determine the storage capacity of the forward pick area along with the SKU assignment and allocation. They also include a constraint on congestion in the forward pick area which ensures that the number of workers per square foot does not exceed a maximum value. This constraint is defined as a function of the time spent in the forward pick area, which in turn depends on the total number of picks and restocks expected. In their solution framework, they solve sub-problems where the capacity of the forward pick area is fixed and objective is to determine the assignment and allocation of SKUs with the capacity and congestion constraints that maximizes the total net benefit of pick savings minus restocking costs. They apply Lagrangian duality arguments to a relaxation of the sub-problem and prove a simple adaptation of the Hackman and Rosenblatt [23] ranking-based heuristic produces a solution that is within the net benefit of one SKU from optimal. Given that typical warehouses have a large number of SKUs, this implies that the solution is near-optimal.

Bartholdi and Hackman [7] also show that the Hackman and Rosenblatt [23] ranking heuristic produces a solution that is near-optimal, though specifically for the formulation (1.3.1). They call the ratio of number of picks to the square root of flow used to rank SKUs *Labor Efficiency*. This ratio and the resulting SKU rank can be calculated a priori. The ranking heuristic offers warehouse managers the insight that SKUs of a higher rank have a higher claim to the forward pick area. This provides an easy way to detect sub-optimal assignments in the forward pick area as changes in demand patterns or SKUs occur. Gu et al. [21] compare the performance of the Hackman and Rosenblatt [23] heuristic solution to the optimal value of (1.3.1) in real warehouse instances. They find that the heuristic solutions are very close to optimal in terms of both the objective value and SKU assignment.

Hackman and Platzman [22] study a form of generalized resource allocation problem where capacity constrained resources need to be allocated to activities. The objective function is assumed to be additively separable but not necessarily concave or differentiable. Each activity uses at most one resource and is restricted to a general set of admissible values. They propose an algorithm using non-smooth convex optimization methods to determine a solution that is near-optimal whenever each allocation is a small fraction of resource capacity. The fluid model for a warehouse with one forward pick area as shown in (1.3.1) and the corresponding extension to multiple forward pick areas is a special case of the resource allocation problem they study. Typical warehouses have many SKUs and so any one SKU occupies only a small portion of any forward pick area. Therefore, the Hackman and Platzman [22] algorithm can be used to solve the assignment and allocation problem to near optimality. Moreover, the restriction of the allocations to a set of admissible values allows consideration of geometries of SKUs and storage units. However, there are two disadvantages of using their algorithm. First, the ranking property is no longer applicable and so we lose a valuable insight on which SKUs deserve to assigned to the forward pick area. Second, this approach requires the implementation of sophisticated non-smooth convex optimization methods.

In an alternative approach to solve the general case of assignment and allocation of SKUs in a warehouse with multiple forward pick areas, Bartholdi and Hackman [6] propose an extension of the Hackman and Rosenblatt [23] ranking-based heuristic. In addition to ranking the SKUs on the basis of labor efficiency, they rank the forward pick areas in descending order of pick savings. Their heuristic is a modification of the Hackman and Rosenblatt [23] heuristic, where the higher ranked forward pick areas are assigned the higher ranked SKUs. They prove that this heuristic produces a near-optimal solution when the number of SKUs is much larger than the number of forward pick areas, as is usually the case. The ranking of the forward pick areas is another valuable insight for warehouse managers that shows the relative claim of forward pick areas over SKUs.

1.4 Other studies on forward pick areas

A related problem is the case of a warehouse with a forward pick area that requires unit-load restocking. Unit-loads require one trip to replenish each unit, e.g. pallets. Bozer [9] was one of the first to study space allocation problems in a pallet rack where the lower level is a forward pick area with unit-load restocking and picking and the higher levels constitute the reserve. Van der Berg et al. [39] analyze how to assign SKUs and allocate space in forward pick areas that store and restock in unit-loads but pick cases from the forward pick area. Their model leads to the insight that there is no need to store more than one unit-load in the forward pick area of a SKU, assuming that restocking can be done instantaneously. Storing more than one unit-load is only for purposes of maintaining safety stock and coordinating the timing of restocking activities. Frazelle et al. [14] and Gu [20] study the problem of determining the optimal size of the forward pick area along with the SKU assignment and allocation. Both papers use the Hackman and Rosenblatt [23] fluid model as a sub-problem within their solution framework to determine SKU assignment and allocation.

1.5 Dissertation overview

In this thesis, we build upon and extend the models of Hackman and Rosenblatt [23] to assign SKUs and allocate space in forward pick areas within a warehouse. The optimal space allocation determined by Hackman and Rosenblatt [23] can lead to arbitrary divisions of space in the forward pick area that may complicate shelf management. In Chapter 2, we construct a more practical allocation where the space allocated to each SKU differs by a power of two and is yet bounded from optimal. There are several advantages to this approach. First, the powers of two allocation accommodates the geometry of the shelf so that allocations are multiples of a standard size (e.g. the width of a lane in a flow rack). Second, as the allocation is a finite set of multiples of a common base, it requires far fewer unique allocation values compared to optimal and resembles the commonly used Equal Space allocation. Third, the restock costs incurred by the powers of two allocation is very close to optimal. Finally, this allocation methodology is simple to implement and can be used within the ranking framework of assigning SKUs to forward pick areas, making it more appealing for warehouse managers.

In Chapter 3, we propose a new proof of the near-optimality of the Hackman and Rosenblatt [23] heuristic to assign and allocate SKUs to a warehouse with one forward pick area. Though the result was already proven by Frazelle et al. [14] using the dual of the formulation (1.3.1), our new argument via the primal formulation, in addition to being simpler, has two benefits. First, given an assignment to the forward pick area that has SKUs ranked lower than one or more SKUs in the reserve, we provide a *direct* way to improve the quality of the assignment and allocation. This is done by *switching* the assigned lower rank SKU with the higher ranked SKU (or SKUs) from the reserve. We explicitly show how to construct such a solution that improves total net benefit while maintaining feasibility. Second, the approach can be generalized to assignment and allocation problems in warehouses with multiple forward pick areas, subject of Chapter 4.

In Chapter 4, we use the new argument developed in Chapter 3 to extend the fluid model to assign and allocate space in warehouses with multiple forward pick areas, and to derive a heuristic that produces a near-optimal solution. Some warehouses are constrained either by limited availability of labor or by layout and equipment used in a manner that limits the number of restocks or picks possible within a given time. We extend the fluid model to account for a cap on number of restocks and picks possible within the planning horizon. We also retain the congestion constraint from Frazelle et al. [14] and show a simple adaption of the Hackman and Rosenblatt [23] heuristic that produces a near-optimal solution in the case of a warehouse with one forward pick area. We generalize this result to the case of the warehouse with multiple forward pick areas with similar set of constraints, again to get a ranking-based algorithm that produces a near-optimal solution.

In the aforementioned three chapters, the basis of allocating space in the forward pick area is the optimal allocations as determined by the formulation in (1.2.1).

However, most warehouses use one of two stocking strategies: Equal Space or Equal Time. In Chapter 5, we show how to determine the optimal assignment of skus to be stocked for a warehouse with one forward pick area that uses one of the two conventional strategies. We also provide an algorithm for the case of a warehouse with multiple forward pick areas that uses either strategy. We show how to implement the algorithms and also test them on data from three different real warehouses. We formulate and analyze the extension of the model to the case of warehouses with constraints on picks and restocks using one of the two conventional strategies.

In Chapter 6, we examine the popularity (picks-per-SKU) distributions from thirty warehouses stocking different classes of items from apparel to service parts for trucks. Warehouses frequently use the 80-20 rule or the ABC distribution or equivalently, the Pareto distribution, to characterize the popularity of SKUs as reflected in the number of requests, or picks. This distribution is used for inventory management decisions within the warehouse. The Pareto distribution, or equivalently, the Power Law, has been identified in many natural processes, from populations of cities to number of visits to websites.

We study whether the data support the assumption of a power law in the popularity distribution. We fit a power law distribution to the data, check if the fitted distribution is indeed suitable, and test other candidate distributions. We find that out of thirty data sets, only sixteen show statistical support for the power law fit. However, out of these sixteen data sets, twelve show a greater support for alternative distributions and only in four data sets, do we see that the power law is the best fit. We test the hypothesis that the picks-per-SKU distributions of warehouses in similar industries are themselves similar. We find that the parameters of the fitted power law distributions vary significantly even within industry categories. Finally, we review the explanations for why power laws arise in other settings and identify explanations that are plausible in the warehouse setting.

CHAPTER II

MANAGEABLE SPACE ALLOCATIONS IN FORWARD PICK AREAS

2.1 Allocating space in a warehouse

The Hackman and Rosenblatt [23] model described in Chapter 1 results in a space allocation that is proportional to the square root of the flow, where flow is the volume of the product forecast to be shipped out during the planning horizon. This means that the allocations can be any positive real number, as the flow can take any value and therefore does not account for SKU and shelf geometries. The result, however, may not be practical when allocating space in a given forward pick area, and the allocations may have to be rounded off to the nearest multiple of the size of an each or carton that can fit in the storage unit. When the size of the each is small, this may not be much of an issue. However, in case of flow racks, the SKU is either stored only as a carton or as eaches in a bin. This implies that the allocation may have to be rounded up or down to the nearest multiple of the respective storage unit (carton or bin), thereby losing the efficiency of the optimal solution. Moreover, rounding down may result in an allocation that does not utilize all the available space in the forward pick area. And rounding up may result in infeasibility due to the total space allocations exceeding the capacity of the forward pick area. Furthermore, the number of distinct numeric values among the allocations could remain high, making management of the forward pick area difficult. To compound the issue, the difficulty of space management increases as new SKUs are introduced and old ones are retired [7].

Another aspect of the optimal allocations is that it results in many unique values of allocations that the warehouse has to manage during the planning horizon. Bartholdi and Hackman [8] show that the optimal allocations share space more evenly than Equal Time. However, given the large number of SKUs in a warehouse, they still could result in too many unique values to be manageable.

Recent work by Walter et al. [40] shows how to determine the optimal allocation when the space can be allocated only in multiples of equally sized units. The allocations are practical if allocating space to SKUs in any multiple of the smallest discrete storage unit is acceptable for the warehouse. In addition, their approach relies on the Hackman and Rosenblatt [23] fluid model and uses the optimal allocations thus generated as a starting solution to round down or up to a multiple of storage unit. Therefore, just like the optimal allocation, this allocation may also result in too many unique values of space allocation to be manageable.

A similar set of issues exists with timing of restocks as well. The optimal allocations could result in each SKU having an arbitrary restocking frequency and therefore cause management overhead when scheduling restocking activities. The Equal Time stocking strategy provides an alternative to warehouse managers that allows them to batch restocking activities. It is, however, not an economical alternative as Bartholdi and Hackman [8] show that the significantly higher restocking cost of Equal Time strategy compared to the optimal outweighs any benefits due to batching.

In this chapter, we present an algorithm to determine space allocations that are easier to manage by being restricted to a constant times integral powers of two, and are still provably close to the optimal restocking cost. We assume that the SKUs to be stored have already been selected, either by the Hackman and Rosenblatt [23] heuristic discussed in Chapter 1, or by another method chosen by the warehouse. Given the SKU assignment, we show how to determine a space allocation of the given form. This allows us to retain the framework of ranking SKUs by labor efficiency to evaluate the claim of a SKU to the forward pick area, a helpful tool for warehouse managers. We construct allocations of the form $\alpha 2^q$, where α is a constant and q

is an integer. These allocations remain feasible and are also space efficient, using the entire available space in the forward pick area. These are easier to manage because the forward pick area is divided into discrete units, multiples of α, and these units differ from each other by a power of 2, resulting in fewer distinct allocation values. This aspect resembles the popularly used Equal Space allocation [8], which is simple to manage because all SKUs have the same allocation. In warehouses that stock sufficiently large number of SKUs, we demonstrate that *almost any* α can be chosen and that the allocation produced is provably near the optimal restocking cost. Therefore, the cost of accommodating the geometries of SKUs and storage is not very significant.

We also show how to extend the same methodology to generate a space allocation such that the restocking intervals are powers of two. This helps to synchronize restocking activities, thereby allowing a warehouse manager to plan and schedule the workforce more easily. As the restocking intervals are restricted to a small finite set, the restocks of SKUs coincide resulting in an allocation that has some of the characteristics of another popularly used strategy, Equal Time [8]. We also show that the space allocation with powers of two restocking intervals is near-optimal in terms of restocking cost.

2.2 Rounding to powers of two

Powers of two approximations have attracted significant attention in the area of order intervals for inventory management problems. Roundy [35] presented an approach to round off optimal order intervals to powers of two for capacitated economic lot sizing problems. The powers of two order intervals are both feasible and provably near-optimal. Powers of two order intervals allow for easier management of items whose reorder points coincide, as the time intervals are constant times powers of two. This approach has been extended to many other settings, including intervals on orders placed by retailers at one warehouse [33], multi-product multi-stage production systems [34], and production and distribution networks [12] among others. In each case, the algorithm produces a solution that maintains feasibility and at the same time is provably near-optimal.

We adapt and extend the powers of two methodology to space allocation in the forward pick areas to generate two types of allocations: powers of two space allocation and powers of two restocking intervals. In the next section we review the Hackman and Rosenblatt fluid model for space allocation, which is the basis of the optimal space allocation. In subsequent sections, we show how to round off the optimal allocations to powers of two space allocations.

2.3 Allocating space using the fluid model

We assume that the SKUs assigned to the forward pick area have already been selected and so the decision to be made is how much space is to be allocated to each SKU. To do so, we use the fluid model of Hackman and Rosenblatt [23] as shown in (1.2.1) in Chapter 1 and restated below.

$$\begin{aligned} \min Z(\mathbf{v}) \equiv \sum_{i \in N} f_i / v_i & \\ \sum_{i \in N} v_i \leq V_F & \qquad\qquad (2.3.1) \\ v_i > 0, & \qquad\qquad i \in N \end{aligned}$$

Following Hackman and Rosenblatt [23], the solution that minimizes total restocks is obtained by applying Lagrangian relaxation to (2.3.1). By doing so we get,

$$\min \sum_{i \in N} f_i / v_i + \lambda \left(\sum_{i \in N} v_i - V_F \right).$$

Setting the first derivative with respect to v_i to zero, we get

$$\lambda = \frac{f_i}{\tilde{v}_i^2}, \qquad (2.3.2)$$

and therefore,

$$\tilde{v}_i = \sqrt{\frac{f_i}{\lambda}}. \tag{2.3.3}$$

Observe that the capacity constraint at optimal allocation has to be tight, as otherwise, space allocated to one of the SKUs can be increased, decreasing the objective and contradicting minimality. Substituting the expression (2.3.3) in constraint (2.3.1), and setting it to equality,

$$\sum_{i \in N} \sqrt{\frac{f_i}{\lambda}} = V_F,$$

the value of λ satisfies the following equation:

$$\sqrt{\lambda} = \frac{1}{V_F} \sum_{i \in N} \sqrt{f_i} V_F. \tag{2.3.4}$$

By substituting (2.3.4) in (2.3.3), we get the optimal solution,

$$\tilde{v}_i = \frac{\sqrt{f_i}}{\sum_{j \in N} \sqrt{f_j}}. \tag{2.3.5}$$

We refer to the space allocation given by expression (2.3.5) as the optimal allocation and denote by $\tilde{\mathbf{v}} = (\tilde{v}_i, i \in N)$. The intermediate expressions derived by analyzing Lagrangian dual will be useful in the development of the powers of two allocations in the following section.

2.4 A powers of two allocation

The optimal allocation can produce space allocations that can be arbitrary real numbers. One way of making the space allocation easier to manage is to round them to a restricted set of harmonious values. We extend the procedure of Roundy [35] to round the optimal allocations to a constant in the interval $[1, 2)$ times an integral power of two. For each SKU i, let r_i denote the integer and z_i the unique value in $[1, 2)$ such that,

$$\tilde{v}_i = z_i 2^{r_i} \qquad 1 \leq z_i < 2. \tag{2.4.1}$$

Let $n = |N|$ and we assume that the set of SKUs $N = \{1, 2, \ldots, n\}$ ordered such that $z_i \leq z_{i+1}$ for $1 \leq i \leq n-1$. We construct a solution of the form $\mathbf{v}^k = (v_i^k, i \in N)$, where

$$v_i^k = \alpha^k 2^{q_i^k} \tag{2.4.2}$$

$$q_i^k = \begin{cases} r_i - 1 & 1 \leq i \leq k \\ r_i & k+1 \leq i \leq n \end{cases} \tag{2.4.3}$$

and α^k is such that it minimizes $Z(\mathbf{v}^k)$. Since the capacity constraint (2.3.1) is tight, we determine α^k to be

$$\alpha^k = \frac{V_F}{\sum_{i \in N} 2^{q_i^k}} \tag{2.4.4}$$

Therefore, the value α^n is

$$\alpha^n = \frac{V_F}{\sum_{i \in N} 2^{r_i - 1}} = \frac{2V_F}{\sum_{i \in N} 2^{r_i}} \leq 2z_n.$$

As $r_i \geq q_i^k$ for all i and k and also that $\sum_{i \in N} \tilde{v}_i = \sum_{i \in N} z_i 2^{r_i} = V$,

$$\alpha^k = \frac{V_F}{\sum_{i \in N} 2^{q_i^k}} > \frac{V_F}{\sum_{i \in N} 2^{r_i}} = \frac{\alpha^n}{2} \geq z_1 \geq 1. \tag{2.4.5}$$

Also note that α^k is increasing in k. Therefore,

$$1 \leq z_1 \leq \frac{\alpha^n}{2} \leq \alpha^1 \leq \alpha^2 \cdots \leq \alpha^n \leq 2z_n < 4. \tag{2.4.6}$$

When $\alpha^k < 2$, the expression (2.4.2) readily gives us a feasible power of two allocation, that satisfies the capacity constraint (2.3.1) and is of the form of a constant in the interval $[1, 2)$ times a power of two. However, it is possible that for some k, $\alpha^k \geq 2$. In such cases, we can rewrite the allocation in the required powers of two form:

$$v_i^k = \frac{\alpha^k}{2} 2^{1+q_i^k}.$$

We adapt Roundy's approach to determining powers of two space allocations for the forward pick area as follows: we determine the powers of two allocations, $\mathbf{v}^k$, for all

Algorithm 1: Algorithm to round off the optimal space allocation to powers of two.

1 Determine z_i and r_i for $i \in N$ using (2.4.1)
2 Index the SKUs as $i \in N = \{1, \dots, n\}$, where $n = |N|$, such that $z_i \leq z_{i+1}$ for $1 \leq i \leq n-1$ and $n = |N|$
3 **for** $k = 1$ **to** n **do**
4   Determine $\mathbf{v}^k$ using (2.4.2), (2.4.3) and (2.4.4) to compute $Z(\mathbf{v}^k)$
5 Select the vector $\mathbf{v}^{\overline{k}}$ that has the least $Z(\mathbf{v}^{\overline{k}})$, $\overline{k} \in 1, \dots, n$

$k = 1, 2, \dots, n$ and choose the one that requires the least number of restocks. The exact method is described in Algorithm 1.

Algorithm 1 runs in $O(|N| \log |N|)$ time since the work required in the algorithm is sorting in Step 2 that takes $O(|N| \log |N|)$, and two $O(|N|)$ loops in Steps 1 and 3 respectively.

Roundy [35] shows that in the case of the single capacity constraint lot sizing problems his procedure obtains a solution that is at most 6% more than the optimal cost. The same bound of 6% holds in our setting as well. That is, the number of restocks for the allocation determined by Algorithm 1 is at most 1.06 times the minimum possible number of restocks. The proof turns out to be a simple adaptation of Roundy's method to this problem, which has a space constraint that is somewhat different from the constraints of the lot sizing problem.

Theorem 2.4.1. The number of restocks in the forward pick area using the powers of two allocation computed in Algorithm 1 is at most 1.06 times the minimum number of restocks.

Following Roundy [35], an upper bound on the powers of two allocation is obtained by bounding the weighted average of the objective computed using the n candidate powers of two solutions in Algorithm 1. The weight assigned to $Z(\mathbf{v}^k)$ is w_k, where

$$w_k = \begin{cases} \log_2(z_{k+1}/z_k) & k < n \\ \log_2(2z_1/z_n) & k = n \end{cases} \tag{2.4.7}$$

Let

$$z_i^k = \begin{cases} z_i & i > k \\ 2z_i & i \le k \end{cases} \tag{2.4.8}$$

Using definitions (2.4.8) and (2.4.3) in (2.4.1), we get

$$\tilde{v}_i = z_i^k 2^{q_i^k}. \tag{2.4.9}$$

Therefore, writing v_i^k in terms of $\tilde{v}_i$ and z_i^k, using (2.4.2), (2.4.3), (2.4.4) and (2.4.9),

$$\begin{aligned} v_i^k &= \alpha^k 2^{q_i^k} \\ &= \left(\frac{V_F}{\sum_{j=1}^n 2^{q_j^k}} \right) \frac{\tilde{v}_i}{z_i^k} \\ &= \left(\frac{V_F}{\sum_{j=1}^n \frac{\tilde{v}_j}{z_j^k}} \right) \frac{\tilde{v}_i}{z_i^k} \end{aligned} \tag{2.4.10}$$

We use the following Lemma from Roundy [35] in the proof:

Lemma 2.4.2. (Roundy) If $i < j$, then

$$\sum_{k=1}^n w_k \left(\frac{z_j^k}{z_i^k} + \frac{z_i^k}{z_j^k} \right) \le (\sqrt{2} + \sqrt{\frac{1}{2}}). \tag{2.4.11}$$

The proof of the theorem is as follows.

Proof. From (2.3.2) we have that,

$$\lambda \tilde{v}_i = \frac{f_i}{\tilde{v}_i} \tag{2.4.12}$$

and

$$Z(\tilde{\mathbf{v}}) = \sum_{i=1}^n \frac{f_i}{\tilde{v}_i} = \sum_{i=1}^n \lambda \tilde{v}_i = \lambda V_F \tag{2.4.13}$$

And hence rewriting the objective using (2.4.10), (2.4.12) and (2.4.13)

$$
\begin{aligned}
Z(\mathbf{v}^k) &= \sum_{i=1}^{n} \frac{\sum_{j=1}^{n} \tilde{v}_j / z_j^k}{V_F} \times \frac{f_i}{\tilde{v}_i / z_i^k} \\
&= \sum_{i=1}^{n} \frac{\sum_{j=1}^{n} \tilde{v}_j / z_j^k}{V_F} \lambda \tilde{v}_i z_i^k \\
&= Z(\tilde{\mathbf{v}}) \sum_{i=1}^{n} \frac{\sum_{j=1}^{n} \tilde{v}_j / z_j^k}{V_F^2} \tilde{v}_i z_i^k \\
&= \frac{Z(\tilde{\mathbf{v}})}{V_F^2} \left\{ \sum_{i=1}^{n} \tilde{v}_i^2 + \sum_{i<j} \tilde{v}_i \tilde{v}_j \left(\frac{z_i^k}{z_j^k} + \frac{z_j^k}{z_i^k} \right) \right\} \qquad (2.4.14)
\end{aligned}
$$

Consider the weighted sum of $Z(\mathbf{v}^k)$ for all k, $1 \leq k \leq n$ with weights w_k, $k = 1, \ldots, n$.

Using the representation of the objective function shown in 2.4.14, we see that

$$
\sum_{k=1}^{n} w_k Z(\mathbf{v}^k) = = \frac{Z(\tilde{\mathbf{v}})}{V_F^2} \left\{ \sum_{i=1}^{n} \tilde{v}_i^2 + \sum_{i<j} \tilde{v}_i \tilde{v}_j \sum_{k=1}^{n} w_k \left(\frac{z_i^k}{z_j^k} + \frac{z_j^k}{z_i^k} \right) \right\}. \qquad (2.4.15)
$$

Using Lemma 2.4.2 and rearranging terms in (2.4.15), we get

$$
\begin{aligned}
\frac{\sum_{k=1}^{n} w_k Z(\mathbf{v}^k)}{Z(\tilde{\mathbf{v}})} &\leq \frac{1}{V_F^2} \left\{ \sum_{i=1}^{n} \tilde{v}_i^2 + \sum_{i<j} \tilde{v}_i \tilde{v}_j \left(\sqrt{2} + \sqrt{0.5} \right) \right\} \\
&\leq \frac{1}{2} \left(\sqrt{2} + \sqrt{0.5} \right) \left(\sum_{i=1}^{n} \tilde{v}_i^2 + \sum_{i<j} 2 \tilde{v}_i \tilde{v}_j \right) / V_F^2 \\
&= \frac{1}{2} \left(\sqrt{2} + \sqrt{0.5} \right) \frac{\left(\sum_{i=1}^{n} \tilde{v}_i \right)^2}{V_F^2} \\
&= \frac{1}{2} \left(\sqrt{2} + \sqrt{0.5} \right) \approx 1.06.
\end{aligned}
$$

We use the fact that the optimal allocations sum up to the space available in the forward pick area, i.e.,

$$
\sum_{i=1}^{n} \tilde{v}_i = V_F.
$$

The number of restocks obtained using the algorithm is

$$
\begin{aligned}
\min_k Z(\mathbf{v}^k) &\leq \sum_{k=1}^{n} w_k Z(\mathbf{v}^k) \\
&\leq 1.06 Z(\tilde{\mathbf{v}}).
\end{aligned}
$$

□

Therefore, the powers of two allocation generated by Algorithm 1 incurs a penalty of at most 6% in additional restocking costs compared to the optimal allocation. In addition, note that the allocation produced by the algorithm is such that it is feasible and uses up the entire space available in the forward pick area. However, to obtain a practical allocation, we need to consider the *manageability* of the powers of two allocation, which can be measured in two ways: the variance of the allocations and the number of distinct values present in the allocation. Warehouse managers consider this variability as important and we address these questions in the the rest of the chapter.

2.5 Manageability of powers of two allocations

From Algorithm 1, we get a space allocation for the SKUs where any two SKUs occupy volumes that differ by a factor of an integral power of two. In addition, the total number of restocks incurred by the space allocation is within 6% of optimal. This allocation also has another property of being *manageable*, with respect to Optimal, Equal Time and Equal Space Allocations. The term manageability is used in the sense of the effort required to manage the allocation: planning, stocking and tracking of the space allocated to SKUs. We analyze the comparisons of manageability among the four methods of allocation in this section.

Given an optimal allocation $\tilde{\mathbf{v}}$, using which we derive the powers of two allocation, the number of possible distinct values for the powers of two allocation is at most $\lfloor \log_2 \max_i \tilde{v}_i \rfloor - \lfloor \log_2 \min_i \tilde{v}_i \rfloor$. This means that there are logarithmically fewer distinct values of space allocated as compared to the optimal allocation. And fewer distinct values of allocations mean that the powers of two space allocation is more manageable for the warehouse.

Another way to analyze manageability, as Bartholdi and Hackman [8] suggest, is to analyze the variance of the space allocations. This measures the evenness of allocations among the SKUs. Bartholdi and Hackman [8] show that the optimal allocation shares labor (or, restocks) among SKUs more evenly than Equal Space allocations and share space more evenly than Equal Time allocations and thereby is more *manageable* than conventional allocation policies. Their analysis is based on comparing the sample variance of space allocations and number of restocks of the optimal allocations with that of Equal Space and Equal Time and showing that they are less than the sample variance of space allocations in Equal Time and sample of variance of number of restocks generated by Equal Space allocations respectively. In this section, we will show a similar result for the powers of two allocations where we compare the variance of the space allocations and number of restocks to the optimal allocations.

To ease notational burdens in the developments to follow, we normalize the value of V_F to 1 and hence, the space allocations as well. Substituting for V_F in equations from (2.4.2) and (2.4.4), we know that the normalized powers of two allocation for SKU i is described as follows:

$$v_i^k = \frac{2^{q_i^k}}{\sum_{j=1}^{n} 2^{q_j^k}}. \tag{2.5.1}$$

Henceforth, the components of the vector $\mathbf{v}^k = (v_1^k, v_2^k, \ldots, v_n^k)$ shall denote these *normalized powers of two allocations*. By construction, of course, the components of $\mathbf{v}^k$ always sum to one.

Since we will be comparing the powers of two allocation to the optimal, we know from Bartholdi and Hackman [8] that the normalized optimal space allocation for SKU i is

$$\frac{\sqrt{f_i}}{\sum_{j=1}^{n} \sqrt{f_j}}.$$

The variance of the space allocated to the SKUs in the powers of two allocation is *close* to that of the optimal allocation and is bounded by a constant factor. To prove this result, we show that the weighted average of the sum of squares of the space allocation as determined in the n solutions generated by the algorithm using the same weights w_k, $1 \leq k \leq n$, defined in (2.4.7) is not that much more than the sum of squares of the optimal space allocation. We use this result to prove that there exists a solution among the n determined whose variance of allocations does not exceed the variance of the optimal beyond a constant factor.

Theorem 2.5.1. There exists some k, $1 \leq k \leq n$, such that the variance of the space allocations, $\mathbf{v}^k$, is at most 1.2637 times the variance of the optimal space allocation.

The following Lemma is useful in the proof.

Lemma 2.5.2. If $i < j$ and $\delta > 0$, then

$$\sum_{k=1}^{n} w_k \left[\delta \left(\frac{z_i^k}{z_j^k} \right)^2 + \left(\frac{z_j^k}{z_i^k} \right)^2 \right] \leq 1.2637(\delta + 1).$$

Proof. Using the definitions (2.4.7) and (2.4.8),

$$\begin{aligned}\sum_{k=1}^{n} w_k \left[\delta \left(\frac{z_j^k}{z_i^k} \right)^2 + \left(\frac{z_i^k}{z_j^k} \right)^2 \right] = &\left(\sum_{1 \leq k < i} \log_2 \frac{z_{k+1}}{z_k} \right) \left[\delta \left(\frac{z_i}{z_j} \right)^2 + \left(\frac{z_j}{z_i} \right)^2 \right] + \\ &\left(\sum_{i \leq k < j} \log_2 \frac{z_{k+1}}{z_k} \right) \left[4\delta \left(\frac{z_i}{z_j} \right)^2 + \frac{1}{4} \left(\frac{z_j}{z_i} \right)^2 \right] + \\ &\left(\sum_{j \leq k < n} \log_2 \frac{z_{k+1}}{z_k} + \log_2 \frac{2z_1}{z_n} \right) \left[\delta \left(\frac{z_i}{z_j} \right)^2 + \left(\frac{z_j}{z_i} \right)^2 \right].\end{aligned}$$

Simplifying, we get

$$\log_2 \frac{2z_i}{z_j} \left[\delta \left(\frac{z_i}{z_j} \right)^2 + \left(\frac{z_j}{z_i} \right)^2 \right] + \log_2 \frac{z_j}{z_i} \left[4\delta \left(\frac{z_i}{z_j} \right)^2 + \frac{1}{4} \left(\frac{z_j}{z_i} \right)^2 \right]. \tag{2.5.2}$$

Setting $x = \frac{z_i}{z_j}$, we rewrite expression (2.5.2) as a function $g(x)$:

$$\begin{aligned} g(x) &\equiv \log_2 2x \left(\delta x^2 + \frac{1}{x^2} \right) + \log_2 \frac{1}{x} \left(4\delta x^2 + \frac{1}{4x^2} \right) \\ &= \delta \left(x^2 \log_2 2x + 4x^2 \log_2 \frac{1}{x} \right) + \left(\frac{1}{x^2} \log_2 2x + \frac{1}{4x^2} \log_2 \frac{1}{x} \right) \\ &\equiv \delta g_1(x) + g_2(x) \end{aligned}$$

where,

$$g_1(x) = x^2 \log_2 2x + 4x^2 \log_2 \frac{1}{x}$$
$$g_2(x) = \frac{1}{x^2} \log_2 2x + \frac{1}{4x^2} \log_2 \frac{1}{x}$$

Since $i < j$, $1 \leq z_i \leq 2$, and z_i is sorted in ascending order, we have that $\frac{1}{2} \leq x \leq 1$. By setting the derivatives of $g_1(x)$ and $g_2(x)$ to zero, we get that for $x \in [0.5, 1]$,

$$\max_{0.5 \leq x \leq 1} g_1(x) = \max_{0.5 \leq x \leq 1} g_2(x) \approx 1.2637.$$

Hence,

$$\max_{0.5 \leq x \leq 1} g(x) \leq \delta \left(\max_{0.5 \leq x \leq 1} g_1(x) \right) + \max_{0.5 \leq x \leq 1} g_2(x) \leq 1.2637(\delta + 1).$$

□

Proof of the theorem is as follows.

Proof. For any allocation $\mathbf{v} = (v_1, \ldots, v_n)$, the sample variance is

$$\sigma_{\mathbf{v}} = \frac{1}{n-1} \left(\sum_{i=1}^{n} v_i^2 - \frac{1}{n} \sum_{i=1}^{n} v_i \right).$$

Since for all allocations the sum of all allocations is V_F, and hence is a constant, we analyze only the first term, the sum of squares of individual volume allocations.

Using (2.5.1), we can write the weighted sum of the sum of squares of allocations of Algorithm 1 as

$$\nu = \sum_{k=1}^{n} \left[w_k \frac{\sum_{i=1}^{n} 2^{2q_i^k}}{\left(\sum_{j=1}^{n} 2^{q_j^k} \right)^2} \right].$$

From (2.4.9) we have that $2^{q_i^k} = \tilde{v}_i / z_i^k$ and hence,

$$\nu = \sum_{k=1}^{n} \left[w_k \frac{\sum_{i=1}^{n} (\tilde{v}_i / z_i^k)^2}{\left(\sum_{j=1}^{n} \tilde{v}_j / z_j^k \right)^2} \right].$$

Since the components of the normalized vector $\tilde{\mathbf{v}}$ sum to unity, we use Jensen's inequality along with the fact that $\sum_k w_k = 1$ to get,

$$\begin{aligned}
\nu &\leq \sum_{k=1}^{n} w_k \sum_{i=1}^{n} \left(\frac{\tilde{v}_i}{z_i^k}\right)^2 \left(\sum_{j=1}^{n} \tilde{v}_j z_j^k\right)^2 \\
&\leq \sum_{k=1}^{n} w_k \sum_{i=1}^{n} \left(\frac{\tilde{v}_i}{z_i^k}\right)^2 \sum_{j=1}^{n} \tilde{v}_j (z_j^k)^2 \\
&\leq \sum_{k} w_k \left\{ \sum_{i} \tilde{v}_i^3 + \sum_{i<j} \tilde{v}_i^2 \tilde{v}_j \left[\frac{\tilde{v}_j}{\tilde{v}_i} \left(\frac{z_i^k}{z_j^k}\right)^2 + \left(\frac{z_j^k}{z_i^k}\right)^2 \right] \right\} \\
&\leq \sum_{i} \tilde{v}_i^3 + \sum_{i<j} \tilde{v}_i^2 \tilde{v}_j \sum_{k} w_k \left[\frac{\tilde{v}_j}{\tilde{v}_i} \left(\frac{z_i^k}{z_j^k}\right)^2 + \left(\frac{z_j^k}{z_i^k}\right)^2 \right]. \qquad (2.5.3)
\end{aligned}$$

Using Lemma 2.5.2 to bound the right hand side in (2.5.3) we have

$$\begin{aligned}
\nu &\leq \sum_{i} \tilde{v}_i^3 + \sum_{i<j} \tilde{v}_i^2 \tilde{v}_j \left[1.2637 \left(\frac{\tilde{v}_j}{\tilde{v}_i} + 1 \right) \right] \\
&\leq 1.2637 \sum_{i} \tilde{v}_i^2 \left(\sum_{j} \tilde{v}_j \right) \\
&\leq 1.2637 \sum_{i} \tilde{v}_i^2.
\end{aligned}$$

This implies there is at least one allocation, $\mathbf{v}^k$, $k \in \{1, \ldots, n\}$, that satisfies the theorem. $\square$

Though the powers of two allocations may add a little more to the variability in allocations, we see that it is still bounded by at most 27% over the corresponding value of the optimal allocation. This bound is meaningful when compared to the variance of the space allocation in Equal Time allocation that many warehouses use in practice. Though, Bartholdi and Hackman [8] show that variance of the optimal allocation is always smaller than the variance of the Equal Time allocation. In practice, however, the variance of the Equal Time space allocation is significantly larger. The reason is that an Equal Time allocation $\mathbf{v}$ has the following property of equal restocks for all SKUs:

$$\frac{f_1}{v_1} = \frac{f_2}{v_2} = \cdots = \frac{f_n}{v_n}$$

This means that the space allocation $\mathbf{v}$ follows the same distribution as the flow of the SKUs. In practice, typically, the distribution of the flow of SKUs in a warehouse has a large support and a heavy tail in the distribution. This means high variance and hence the space allocation also has a proportionally large variance as well. On the other hand, the optimal allocation is proportional to the square root of flow, which has a significantly smaller variance than the flow as the square root transformation compresses the upper end of the distribution. In an example we consider later in the chapter, the variance of the Equal Time allocation is almost 9 times as large as the variance of the optimal allocation. Given this big difference between the optimal and Equal Time allocations, the variance of the powers of two allocation not being more than 27% than the optimal is a meaningful bound.

The number of restocks for each SKU is the indicator of labor involved in maintaining the forward pick area. Bartholdi and Hackman [8] show that under optimal allocations, SKU i incurs a fraction of total restocks equal to $\sqrt{f_i}/\sum_j \sqrt{f_j}$, which is the same as the fraction of space occupied by the SKU, $\tilde{v}_i$. The variance of the number of restocks for each SKU obtained from powers of two allocations is also bounded from the variance of the optimal number of restocks for each SKU in a similar manner to the space allocations. We show in the following theorem, using an approach very similar to Theorem 2.5.1, that for at least one of the n solutions generated by Algorithm 1, the variance of the number of restocks is at most 1.2637 times the number of restocks generated by the optimal allocation.

Theorem 2.5.3. There exists some k, $1 \leq k \leq n$, such that the variance of the number of restocks generated by the k^{th} solution from Algorithm 1 is at most 1.2637 times the variance of the number of restocks generated by the optimal allocation.

Proof. For any allocation $\mathbf{v} = (v_1, \ldots, v_n)$, the number of restocks of a SKU i is f_i/v_i

and hence the sample variance, σ_r^2, of the number of restocks is

$$\sigma_r^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n}\left(\frac{f_i}{v_i}\right)^2 - \frac{1}{n}\sum_{i=1}^{n}\frac{f_i}{v_i}\right].$$

For any allocation, the sum of the restocks is at least as much as the optimal number of restocks. And, since we are interested in an upper bound for σ_r^2 in comparison to the corresponding value for the optimal allocation, we can ignore the term $\sum_{i=1}^{n} f_i/v_i$. We need only bound the sum of squares of the number of restocks with respect to the sum of squares of number of restocks generated by the optimal allocation. And so, if we show that

$$\sum_{i=1}^{n}\left(\frac{f_i}{v_i}\right)^2 \leq 1.2637\sum_{i=1}^{n}\left(\frac{f_i}{\tilde{v}_i}\right)^2$$

then,

$$\begin{aligned}\sigma_r^2 &= \frac{1}{n-1}\left[\sum_{i=1}^{n}\left(\frac{f_i}{v_i}\right)^2 - \frac{1}{n}\sum_{i=1}^{n}\frac{f_i}{v_i}\right]\\ &\leq \frac{1}{n-1}\left[\sum_{i=1}^{n}\left(\frac{f_i}{\tilde{v}_i}\right)^2 - \frac{1}{n}\sum_{i=1}^{n}\frac{f_i}{v_i}\right].\end{aligned}$$

In the case of the powers of two allocation, the number of restocks of SKU i is determined as f_i/v_i^k. From (2.5.1) and using (2.4.9) to get $2^{q_i^k} = \tilde{v}_i/z_i^k$, we have that,

$$\frac{f_i}{v_i^k} = \frac{f_i}{2^{q_i^k}}\sum_{j=1}^{n} 2^{q_j^k} = \frac{f_i}{\tilde{v}_i/z_i^k}\sum_{j=1}^{n}\tilde{v}_j/z_j^k.$$

From (2.3.2), we have that $\lambda\tilde{v}_i = f_i/\tilde{v}_i$ and observing that λ does not depend on i, the proportion of restocks of SKU i is

$$\begin{aligned}\frac{f_i}{\tilde{v}_i/z_i^k}\sum_{j=1}^{n}\tilde{v}_j/z_j^k &= \lambda\tilde{v}_i z_i^k\sum_{j=1}^{n}\tilde{v}_j/z_j^k\\ &\propto \tilde{v}_i z_i^k\sum_{j=1}^{n}\tilde{v}_j/z_j^k.\end{aligned}$$

Hence the fraction of total restocks of SKU i is,

$$\frac{\tilde{v}_i z_i^k}{\sum_{j=1}^{n}\tilde{v}_j z_j^k}.$$

The weighted sum of sum of squares of the number of restocks is

$$\tau = \sum_{k=1}^{n} w_k . \sum_{i=1}^{n} \frac{(\tilde{v}_i z_i^k)^2}{(\sum_{j=1}^{n} \tilde{v}_j z_j^k)^2}.$$

Note that the normalized components of $\tilde{\mathbf{v}}$ sum to unity. Therefore, using Jensen's inequality and rearranging terms, we get

$$\begin{aligned}
\tau &\leq \sum_{k=1}^{n} w_k \sum_{i=1}^{n} (\tilde{v}_i z_i^k)^2 \left(\sum_{j=1}^{n} \tilde{v}_j / z_j^k \right)^2 \\
&\leq \sum_{k=1}^{n} w_k \sum_{i=1}^{n} (\tilde{v}_i z_i^k)^2 \sum_{j=1}^{n} \tilde{v}_j / (z_j^k)^2 \\
&\leq \sum_{k} w_k \left\{ \sum_{i} \tilde{v}_i^3 + \sum_{i<j} \tilde{v}_i^2 \tilde{v}_j \left[\frac{\tilde{v}_j}{\tilde{v}_i} \left(\frac{z_j^k}{z_i^k} \right)^2 + \left(\frac{z_i^k}{z_j^k} \right)^2 \right] \right\} \\
&\leq \sum_{i} \tilde{v}_i^3 + \sum_{i<j} \tilde{v}_i^2 \tilde{v}_j \sum_{k} w_k \left[\frac{\tilde{v}_j}{\tilde{v}_i} \left(\frac{z_j^k}{z_i^k} \right)^2 + \left(\frac{z_i^k}{z_j^k} \right)^2 \right].
\end{aligned} \tag{2.5.4}$$

Using Lemma 2.5.2 to bound the right hand side in (2.5.4) we have

$$\begin{aligned}
\tau &\leq \sum_{i} \tilde{v}_i^3 + \sum_{i<j} \tilde{v}_i^2 \tilde{v}_j \left[1.2637 \left(\frac{\tilde{v}_j}{\tilde{v}_i} + 1 \right) \right] \\
&\leq 1.2637 \sum_{i} \tilde{v}_i^2 \left(\sum_{j} \tilde{v}_j \right) \\
&\leq 1.2637 \sum_{i} \tilde{v}_i^2.
\end{aligned} \tag{2.5.5}$$

As noted earlier, from Bartholdi and Hackman [8], under optimal allocations, the fraction of space allocated to SKU i is the same as the fraction of total restocks due to the SKU. Hence, in (2.5.5), the weighted sum of sum of squares of normalized powers of two allocations is less than 1.2637 times sum of squares of fractions of total restocks generated by each SKU. This implies there is at least one allocation, $\mathbf{v}^k$, $k \in \{1, \ldots, n\}$, that satisfies the Theorem. □

Like in the case of the variance of the space allocation, we compare the variance of number of restocks generated by the powers of two allocation to the variance of number of restocks of another commonly used allocation strategy, the Equal Space allocation. The number of restocks in an Equal Space allocation is proportional to the flow, whereas the number of restocks for the optimal allocation is proportional to the square root of flow. For the same reasons as in the case of space allocations, this means that in practice, the variance of number of restocks in the Equal Space allocation is typically much greater than the corresponding value for the optimal allocation. Therefore, the bound of the variance in number of restocks of the powers of two allocation not exceeding the corresponding value for the optimal allocation by not more than 27% is a useful one.

Theorems 2.5.1 and 2.5.3 show that there exists a powers of two allocation that has a variance not much greater than the optimal. However, this solution need not be the best powers of two solution that guarantees to be within 6% of optimal as we would need to consider other candidate solutions among the n generated by Algorithm 1 if we need to bound our solution based on their variance. Therefore, we need to understand how much more expensive can this chosen solution be. The following theorem shows that even the most expensive of the n solutions generated by the powers of two allocation algorithm requires at most 12.5% more restocks than the optimal allocation.

Theorem 2.5.4. Each of the solutions, $\mathbf{v}^k$, $1 \leq k \leq n$, produced by Algorithm 1 generates at most 12.5% restocks more than the optimal allocation.

To prove the theorem we will need the following lemma.

Lemma 2.5.5. The objective value of the following is at most $1.125(\sum_{i=1}^{n} \tilde{v}_i)^2$.

$$\max \sum_{i=1}^{n} \tilde{v}_i^2 + \sum_{i<j} \tilde{v}_i \tilde{v}_j \left(\frac{z_i}{z_j} + \frac{z_j}{z_i} \right)$$

$$\text{s.t } 1 \leq z_1 \leq z_2 \leq \cdots \leq z_n \leq 2$$

Proof. Extend the vector $\mathbf{z}$ by adding two elements $z_0 = 1$ and $z_{n+1} = 2$. The optimal solution, $\tilde{\mathbf{z}}$, to the math program lies on the boundary and for some l, $1 \leq l \leq n+1$,

$\tilde{z}_i = 1$ for all $i < l$, and $\tilde{z}_i = 2$ for all $i \geq l$. Suppose not. Then the optimal solution $\tilde{\mathbf{z}}$ is such that for there exists u and v such that $\tilde{z}_u > \tilde{z}_{u-1}$, $\tilde{z}_v < \tilde{z}_{v+1}$ and $v \geq u$. We choose least such v. Hence, it must be that

$$\tilde{z}_u = \tilde{z}_{u+1} = \cdots = \tilde{z}_v.$$

Consider the reformulated math program where $\mathbf{z}$ is fixed at the optimal solution except the u^{th} to the v^{th} components, which are set as equal to, say, z. As z is the only variable, isolating the terms involving z, we get

$$\max \sum_{u \leq i \leq v, j < u} \tilde{v}_i \tilde{v}_j \left(\frac{z}{\tilde{z}_j} + \frac{\tilde{z}_j}{z} \right) + \sum_{u \leq i \leq v, j > v} \tilde{v}_i \tilde{v}_j \left(\frac{z}{\tilde{z}_j} + \frac{\tilde{z}_j}{z} \right)$$

$$\text{s.t } \tilde{z}_{u-1} \leq z \leq \tilde{z}_{v+1}$$

The second derivative of the objective function is

$$\frac{1}{z^2} \left(\sum_{u \leq i \leq v, j < u} \tilde{v}_i \tilde{v}_j \tilde{z}_j + \sum_{u \leq i \leq v, j > v} \tilde{v}_i \tilde{v}_j \tilde{z}_j \right) > 0,$$

and hence the function is strictly convex. The maximum value of a strictly convex function is at its boundary and hence it occurs at $z = \tilde{z}_{u-1}$ or $z = \tilde{z}_{v+1}$. But this contradicts our assumption that the optimal solution for z is $\tilde{z}_u$, that is, in the interior of the interval $[\tilde{z}_{u-1}, \tilde{z}_{v+1}]$.

Therefore, the optimal solution is of the form $\tilde{z}_i = 1$ for all $i < l$ and $\tilde{z}_i = 2$ for all $i \geq l$. Rewriting the objective function, we get

$$
\begin{aligned}
&\sum_{i=1}^{n} \tilde{v}_i^2 + \sum_{i<j<l} 2\tilde{v}_i\tilde{v}_j + \sum_{l\leq i<j} 2\tilde{v}_i\tilde{v}_j + \sum_{i<l,j\geq l} \frac{5}{2}\tilde{v}_i\tilde{v}_j \\
&= \left(\sum_{i=1}^{n} \tilde{v}_i\right)^2 + \frac{1}{2}\sum_{i<l,j\geq l} \tilde{v}_i\tilde{v}_j \\
&= \left(\sum_{i=1}^{n} \tilde{v}_i\right)^2 + \frac{1}{2}\left(\sum_{i<l} \tilde{v}_i\right)\left(\sum_{i\geq l} \tilde{v}_i\right) \\
&= \left(\sum_{i=1}^{n} \tilde{v}_i\right)^2 + \frac{1}{2}\left(\sum_{i<l} \tilde{v}_i\right)\left(\sum_{i} \tilde{v}_i - \sum_{i<l} \tilde{v}_i\right) \\
&\leq \left(\sum_{i=1}^{n} \tilde{v}_i\right)^2 + \frac{1}{2}\left(\frac{\sum_i \tilde{v}_i}{2}\right)^2 \\
&\leq \frac{9}{8}\left(\sum_{i=1}^{n} \tilde{v}_i\right)^2 = 1.125\left(\sum_{i=1}^{n} \tilde{v}_i\right)^2,
\end{aligned}
$$

using the fact that the maximum of $x(C-x)$ such that $0 \leq x \leq C$ is $C^2/4$, where, $x = \sum_{i<l} \tilde{v}_i$ and $C = \sum_i \tilde{v}_i$. □

The proof of the theorem is as follows.

Proof. We observe from equation (2.4.14) that,

$$
Z(\mathbf{v}^k) = \frac{Z(\tilde{\mathbf{v}})}{V_F^2}\left[\sum_{i=1}^{n} \tilde{v}_i^2 + \sum_{i<j} \tilde{v}_i\tilde{v}_j\left(\frac{z_i^k}{z_j^k} + \frac{z_j^k}{z_i^k}\right)\right].
$$

Since the optimal allocation vector is normalized,

$$
\frac{Z(\mathbf{v}^k)}{Z(\tilde{\mathbf{v}})} = \sum_{i=1}^{n} \tilde{v}_i^2 + \sum_{i<j} \tilde{v}_i\tilde{v}_j\left(\frac{z_i^k}{z_j^k} + \frac{z_j^k}{z_i^k}\right). \tag{2.5.6}
$$

We note that among the solutions, $Z(\mathbf{v}^k)$ for $1 \leq k \leq n$, generated by Algorithm 1, the only variant is $\mathbf{z}^k$. Using the ordering of $\mathbf{z}$ in step (2) of Algorithm 1, and the definition of $\mathbf{z}^k$ in equation (2.4.8), we observe that

$$
\begin{aligned}
& 1 \leq z_1 \leq z_2 \leq \cdots \leq z_n \leq 2 \\
\Rightarrow \quad & 1 \leq z_{k+1} \leq \cdots \leq z_n \leq 2 \leq 2z_1 \leq \cdots \leq 2z_k \leq 2z_{k+1} \\
\Rightarrow \quad & z_{k+1}^k \leq \cdots \leq z_n^k \leq z_1^k \leq \cdots \leq z_k^k \leq 2z_{k+1},
\end{aligned} \tag{2.5.7}
$$

and hence,

$$\frac{1}{2} \leq \frac{z_i^k}{z_j^k} \leq 2 \qquad\qquad \forall i, j, k \in \{1, \ldots, n\}. \tag{2.5.8}$$

We observe that in the restocking cost of the k^{th} solution, given in Expression (2.5.6), $\mathbf{z}$ appears only as a ratio of two of its components and in Expression (2.5.8), that these ratios are at least 0.5 and at most 2. Therefore, for any given k, we can scale $\mathbf{z}^k$ by by a suitable factor to get a scaled vector $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ so that all values lie between 1 and 2 and the ratio z_i^k / z_j^k is the same as the corresponding ratio y_i / y_j for the scaled vector. We estimate an upper bound for $\frac{Z(\mathbf{v}^k)}{Z(\tilde{\mathbf{v}})}$ using the following math program.

$$\max \sum_{i=1}^{n} \tilde{v}_i^2 + \sum_{i<j} \tilde{v}_i \tilde{v}_j \left(\frac{y_i}{y_j} + \frac{y_j}{y_i} \right)$$

$$\text{s.t } 1 \leq y_{k+1} \leq y_{k+2} \leq \cdots \leq y_n \leq y_1 \leq \cdots \leq y_k \leq 2$$

From Lemma 2.5.5, the maximum value of the above math program is at most $1.125(\sum_i \tilde{v}_i)^2 = 1.125$ since by normalization $\sum_i \tilde{v}_i = 1$. Hence, $Z(\mathbf{v}^k) \leq 1.125 Z(\tilde{\mathbf{v}})$ for $1 \leq k \leq n$. Therefore the solutions produced by Algorithm 1 generate at most 12.5% more restocks than the optimum. $\square$

Therefore, we would incur an additional restocking cost of at most 12.5% regardless of which allocation we choose from the set of possible $|N|$ solutions. At times, it is entirely possible that the least cost solution may not be practical not just for purposes of variance, or, too many distinct space allocations to manage. The allocation can be impractical because of the constant in the powers of two form as the type of shelving may not allow for such an allocation or multiples thereof. This result means that if among the n solutions, the one with the most practical value of α^k is chosen such that the shelf multiples are realistic, we incur no more than 12.5% restocks over the optimal. This allows the warehouse manager to choose from among the remaining solutions with a penalty that is bounded.

The algorithm to determine the powers of two allocations generates only n distinct values and the warehouse manager has to choose the one that is most practical. For example, if the only feasible space allocation in the given forward pick area is in multiples of 1.5 feet, then only the allocations of the form 1.5 times a power of two will be suitable. If among the $|N|$ solutions produced by Algorithm 1, there is a powers of two allocation with a constant term that is 1.5 or close enough to 1.5, then we have the desired allocation. And by Theorem 2.5.4, this allocation incurs no more than an additional 12.5% in restocking cost over the optimal. But, does a solution always exist? We analyze this question in the next section.

2.6 Choosing the constant

The shelf space allocated to SKU i by solution k is $v_i^k = \alpha^k 2^{q_i^k}$. Therefore, the shelf spaces allocated are $\ldots, \frac{1}{8}, \frac{1}{4}, \frac{1}{2}, 1, 2, 4, 8, \ldots$ times a fixed constant. The powers of two algorithm also generates as many solutions as there are SKUs. Warehouses, particularly in North America, typically stock thousands of SKUs. Therefore, a typical shelf space allocation using Algorithm 1 may generate thousands of solutions. We know from Theorem 2.5.4 that every one of those solutions is within 12.5% of optimal and from Theorem 2.4.1 that there is at least one solution that is guaranteed to be within 6% of optimal. If the solution with the least restocking cost has a constant that is unsuitable to the way the shelf space may be divided at the warehouse under consideration, then we could potentially choose among the other solutions, if one of them has a suitable constant. Does a powers of two allocation whose constant may be considered *suitable* to the shelving preferences of the warehouse always exist? We answer this question by looking at the *density* of the constant values generated by the n powers of two allocations.

Algorithm 1 generates n solutions, each with a constant term α^k, and according to (2.4.4) they are determined as follows:

$$\alpha^k = \frac{V_F}{\sum_{i=1}^n 2^{q_i^k}}.$$

We add a notional value corresponding to $k = 0$, α^0, defined as

$$\alpha^0 = \frac{V_F}{\sum_{i=1}^n 2^{r_i}}. \tag{2.6.1}$$

Note that the powers of two allocation corresponding to $k = 0$ is identical to $k = n$. As $q_i^n = r_i - 1$ for $i \in \{1, \ldots, n\}$, we see that $\alpha^n = 2\alpha^0$. From (2.4.6), we know that $\alpha^k \in (\alpha^0, 2\alpha^0]$ for $k = 1, 2, \ldots, n$. The values of α^k that lie in the interval $(\alpha^0, 2)$ is the same as the constant term in the powers of two form of the corresponding allocations. However, when $\alpha^k \in [2, 2\alpha^0]$, we divide the value by two to get the resulting constant, which in turn lies in the $[1, \alpha^0]$ in the corresponding allocations. Therefore, the range of the constant in the powers of two form is lies in the interval $[1, 2)$. And they are at least as dense in the interval as the values of α^k in the interval $(\alpha^0, 2\alpha^0]$. In Theorem 2.6.1, we analyze the density of the values α^k and show that in the the limit as the number of SKUs goes to infinity, the difference between α^k and α^{k+1} converges to zero.

Theorem 2.6.1. If the powers of two, $\{2^{r_i}\}_n$ are independent and identically distributed with finite mean and variance, then the difference between any two consecutive values of α^k converges to zero, i.e., for all k, as $n \to \infty$,

$$\alpha^k - \alpha^{k-1} \to 0.$$

Further, this convergence occurs at a rate of at least $n\sqrt{n}$.

Proof. Let $\{2^{r_i}\}_{i=1,\ldots,n}$ be distributed according to a cumulative distribution function F with mean μ and standard deviation σ. Let $\{X_i\}_{i=1,\ldots,n}$ be a sample from the distribution. Consider the following statistic:

$$S_n^k = \frac{1}{2}\sum_{i=1}^{k} X_i + \sum_{i=k+1}^{n} X_i.$$

The mean μ_n^k and sample standard deviation s_n^k of S_n^k is

$$\mu_n^k = \left(n - \frac{k}{2}\right)\mu$$
$$s_n^k = \sigma\sqrt{(n - \frac{3}{4}k)}$$

For notational convenience, we omit the superscript k. By Lindeberg's Central Limit Theorem, if the *Lindeberg's condition* shown in (2.6.2) holds true, then $\frac{S_n - \mu_n}{s_n}$ converges in distribution to the standard normal variable as $n \to \infty$.

$$\forall \varepsilon > 0,\ \lim_{n\to\infty} \frac{1}{s_n^2} \sum_{i=1}^{n} \int_{|x-\nu_i|>\varepsilon s_n} (x - \nu_i)^2 dF_i(x) = 0 \tag{2.6.2}$$

ν_i and F_i are the mean and distribution function of $\frac{1}{2}X_i$ if $1 \le i \le k$ and X_i if $k < i \le n$. Since s_n increases without bound as $n \to \infty$ (knowing $k < n$), the measure of the domain over which the integration is performed tends to 0 and the denominator goes to infinity. Hence the condition (2.6.2) holds true and

$$\frac{S_n - \mu_n}{s_n} \xrightarrow{D} N(0,1).$$

Substituting for s_n and μ_n and dividing by n, we get

$$\sqrt{n}\left[\frac{\frac{S_n}{n} - \left(1 - \frac{k}{2n}\right)\mu}{\sigma\sqrt{1 - \frac{3k}{4n}}}\right] \xrightarrow{D} N(0,1).$$

Since $k \le n$, $\lim_{n\to\infty} \frac{k}{n} = \gamma$ where $\gamma \in [0,1]$. By laws of limits we get,

$$\sqrt{n}\left[\frac{\frac{S_n}{n} - \left(1 - \frac{1}{2}\gamma\right)\mu}{\sigma\sqrt{1 - \frac{3}{4}\gamma}}\right] \xrightarrow{D} N(0,1). \tag{2.6.3}$$

Consider the statistic, $T_n^k = S_n^k/n$. From (2.6.3), for any k, as $n \to \infty$, T_n^k and T_n^{k+1} converge to the same normal distribution as $\frac{k}{n}$ and $\frac{k+1}{n}$ converge to the same value γ. Let $\bar{\mu} = \left(1 - \frac{1}{2}\gamma\right)\mu$ and $\bar{\sigma} = \sigma\sqrt{1 - \frac{3}{4}\gamma}$. And hence

$$\sqrt{n}\left(T_n^k - \bar{\mu}\right) \xrightarrow{D} N(0, \bar{\sigma}^2)$$
$$\sqrt{n}\left(T_n^{k+1} - \bar{\mu}\right) \xrightarrow{D} N(0, \bar{\sigma}^2)$$

We rewrite the two random variables as a sequence of random 2-vectors, $\{\mathbf{T}_n\}$, as follows:

$$\mathbf{T}_n = \begin{bmatrix} T_n^k \\ T_n^{k+1} \end{bmatrix}.$$

Since T_n^k and T_n^{k+1} converge to the same normal distribution, the mean $\bar{\mu}$ and covariance matrix θ of $\{\mathbf{T}_n\}$ are

$$\theta = \begin{bmatrix} \bar{\mu} \\ \bar{\mu} \end{bmatrix}$$

$$\varSigma = \begin{bmatrix} \sigma & \sigma \\ \sigma & \sigma \end{bmatrix}$$

Hence, we have,

$$\sqrt{n}\,(\mathbf{T}_n - \theta) \xrightarrow{D} N(0, \varSigma). \tag{2.6.4}$$

We define a function, $g : \Re^2 \to \Re$,

$$g(\mathbf{x}) = \frac{1}{x_1} - \frac{1}{x_2}.$$

and whose gradient, ∇g, is

$$\nabla g(\mathbf{x}) = \begin{bmatrix} \frac{-1}{x_1^2} \\ \frac{1}{x_2^2} \end{bmatrix}.$$

Therefore,

$$g(\mathbf{T}_n) = \frac{1}{T_n^k} - \frac{1}{T_n^{k+1}}$$

and

$$\nabla g(\theta) = \begin{bmatrix} \frac{-1}{\bar{\mu}^2} \\ \frac{1}{\bar{\mu}^2} \end{bmatrix}. \tag{2.6.5}$$

Given (2.6.5), we see that $\nabla g(\theta)$ is non-null and continuous in the neighborhood of θ since $\bar{\mu} > 0$. Using this observation along with convergence of $\mathbf{T}_n$ in (2.6.4), we use Theorem 3.4.5 in Sen and Singer [36] to obtain the following convergence for $g(\mathbf{T}_n)$:

$$\sqrt{n}\{g(\mathbf{T}_n) - g(\theta)\} \xrightarrow{D} N(0, [\nabla g(\theta)]^T \Sigma \nabla g(\theta)). \tag{2.6.6}$$

Estimating the mean and variance in (2.6.6), we get

$$g(\theta) = \frac{1}{\bar{\mu}} - \frac{1}{\bar{\mu}} = 0$$

$$[\nabla g(\theta)]^T \Sigma \nabla g(\theta) = \begin{bmatrix} \frac{-1}{\bar{\mu}^2} & \frac{1}{\bar{\mu}^2} \end{bmatrix} \begin{bmatrix} \sigma & \sigma \\ \sigma & \sigma \end{bmatrix} \begin{bmatrix} \frac{-1}{\bar{\mu}^2} \\ \frac{1}{\bar{\mu}^2} \end{bmatrix} = 0$$

Hence, simplifying (2.6.6), we get

$$\sqrt{n}\left(\frac{n}{S_n^k} - \frac{n}{S_n^{k+1}}\right) \xrightarrow{D} N(0,0) \equiv 0,$$

which is equivalent to,

$$n\sqrt{n}\left(\frac{1}{S_n^k} - \frac{1}{S_n^{k+1}}\right) \xrightarrow{P} 0.$$

And hence the convergence

$$\frac{1}{S_n^k} - \frac{1}{S_n^{k+1}} \xrightarrow{P} 0$$

happens at least at the rate of $n\sqrt{n}$. Now, the difference between the values of α^k and α^{k+1} is,

$$\alpha^k - \alpha^{k+1} = V_F\left(\frac{1}{S_n^k} - \frac{1}{S_n^{k+1}}\right) \xrightarrow{P} 0$$

which converges at the rate of $n\sqrt{n}$. □

The result shows that as n becomes large, the set of values α^k become very dense in the interval $(\alpha^0, 2\alpha^0]$ with the difference between any two adjacent values converging to zero at a fast rate. Therefore the corresponding constant values normalized to the It immediately follows that when the number of SKUs is large, which is almost always the case in warehouses, the warehouse manager can choose a powers of two solution with almost any constant between $[1, 2)$. In addition, we know that the cost of choosing the solution is not significant as we have already shown that this solution is adds no more than an additional 12.5% restocks over the optimal. In the next section we look at an example data set from a warehouse and evaluate the quality of the powers of two allocations generated by Algorithm 1.

2.7 An example

We use Algorithm 1 to calculate the powers of two allocation for a data set from a warehouse of the telecommunications provider studied in Bartholdi and Hackman [8]. The warehouse has 3050 SKUs and we use the flows of the SKUs in a given time period to determine the allocations. The algorithm generates 3050 powers of two allocations and the best powers of two solution is within 3.9% of optimal and the worst, 4.2% of optimal, making them almost nearly identical in cost. As a comparison, the Equal Time or Equal Space allocations cost 97% more than the optimal, which is significantly more than the powers of two allocations. In addition, adjacent values of α^k differs by at most 0.005, giving us an extensive choice of the powers of two allocation with a suitable constant term. That is, given any suitable value in interval $[1, 2)$, a corresponding powers of two allocation can be chosen with a constant term that is within 0.005 of the chosen value. Among the 3050 allocations generated by the algorithm, the variance of the space allocated to the SKUs range from 3% to 17% more than the variance of optimal allocations. Equal Time, in comparison, produced an allocation that had a variance of 890% more than the variance for optimal allocations. In terms of number of restocks, the variance of the powers of two solutions ranges between 8% and 23% more than optimal, where as, the Equal Space allocation, in comparison, has a very high variance of 3730% more than the optimal solution. We see that the solution quality of the powers of two algorithm retains many of the useful properties as shown in Bartholdi and Hackman [8] with respect to optimality and manageability and at the same time brings in a useful aspect of Equal Space allocations, where the allocations though not identical, differ only by multiples of two and have only a few unique values.

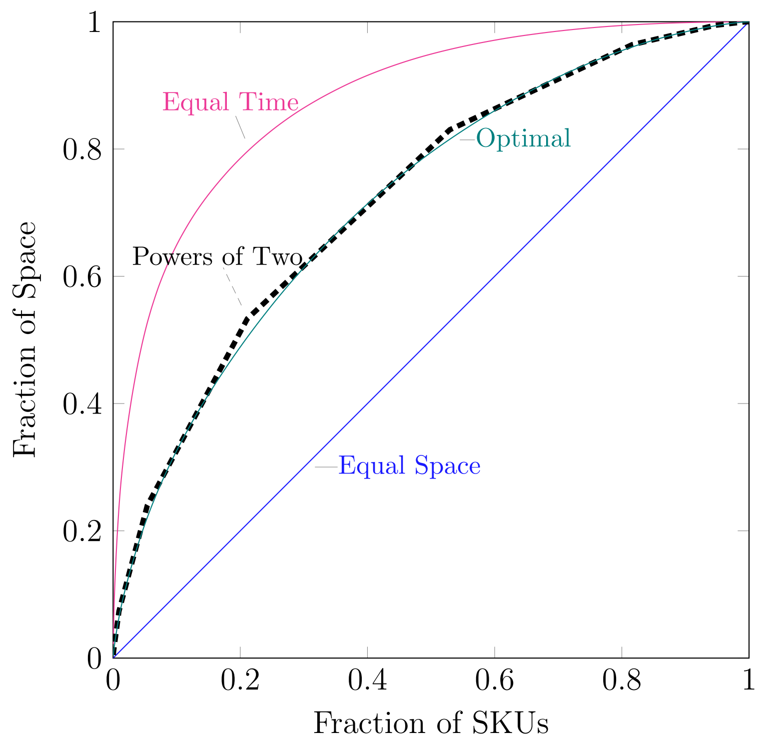

Figure 4: Comparison of space allocation of Equal Time, Powers of Two, Optimal and Equal Space schemes. The Powers of Two allocation shares space among SKUs almost as evenly as the Optimal and both allocations share space much more evenly than Equal Time.

Figures 4 and 5 show that the distribution of powers of two space allocations is very similar to the optimal allocation. Figure 4 shows that the powers of two allocation, just like the optimal allocation, is in between the Equal Space and Equal Time allocations. The powers of two allocation has a more even space distribution than Equal Time, almost as even as the optimal allocation, and at the same time, with significantly lower restock cost than Equal Time or Space allocations.

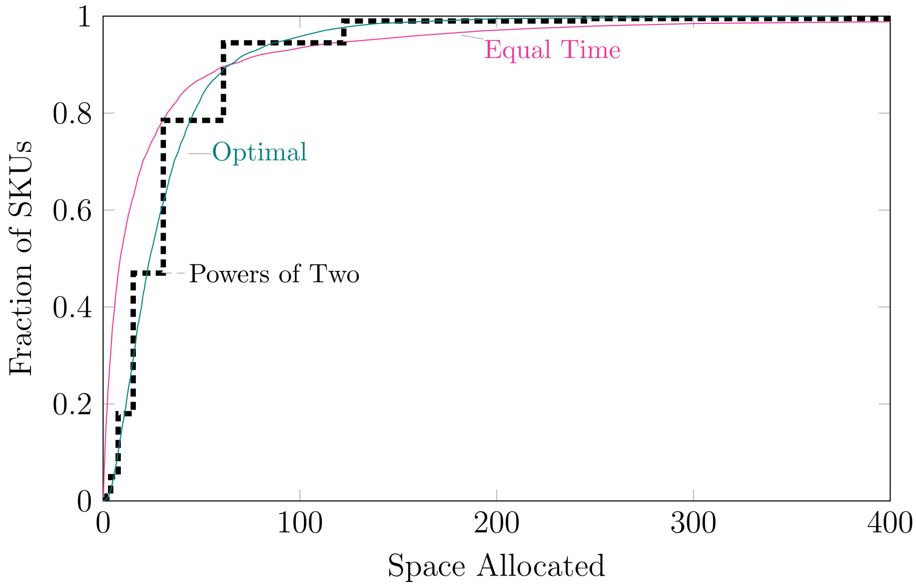

Figure 5: Cumulative distribution function of space allocations of Equal Time, Optimal and Powers of Two schemes. Observe that the Powers of Two requires only a few distinct values of space allocation.

When comparing the number of restocks between the allocation schemes in Figure 6 and Figure 7, we see a similar behavior. The powers of two and optimal allocations are very similar, and share the number of restocks almost as evenly among the SKUs. However, both allocations share restocks among the SKUs much more evenly than Equal Space.

Figures 8 and 9 show histograms that compare the space allocations and number of restocks of the powers of two scheme to the others. We see from Figure 8 that the powers of two scheme, just like the optimal allocation [8], has less variability compared to Equal Space (with respect to number of restocks) and Equal Time (with respect to space allocations). Figure 9 shows that the variability of space allocations and number of restocks for powers of two scheme is very similar to that of the optimal allocation.

The example demonstrates that the powers of two allocation is very effective when it comes to manageability in terms of units of space allocated, very few distinct values of space allocated with the added advantage of being close to the optimal in terms of

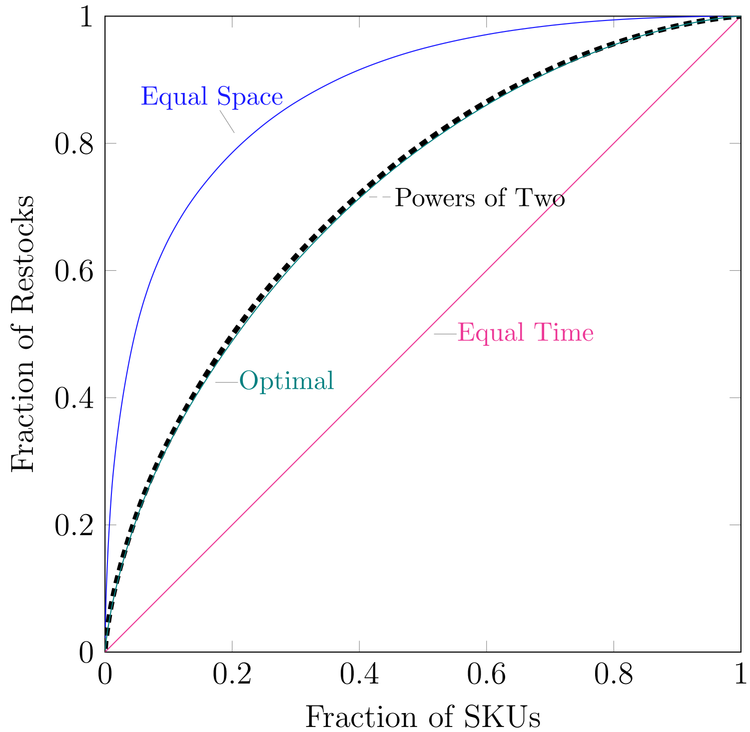

Figure 6: Comparison of number of restocks required by Equal Time, Optimal and Powers of Two schemes. The number of restocks for the powers of two allocation is almost as evenly distributed as the optimal and significantly more so than for the Equal Space allocation.

number of restocks as well as variability in space allocated and number of restocks of SKUs.

2.8 Powers of two restocking intervals

Just like the powers of two space allocations, the optimal restocking time intervals can also be rounded off to a constant factor times a power of 2. This type of allocation is similar to the Equal Time allocation, where we synchronize restocking intervals. We show that type of allocation can be done without a significant increase in restocking cost as compared to the optimal solution. This is similar to Roundy [35], where the order time intervals for the lot sizing problem are set to powers of two so that they coincide.

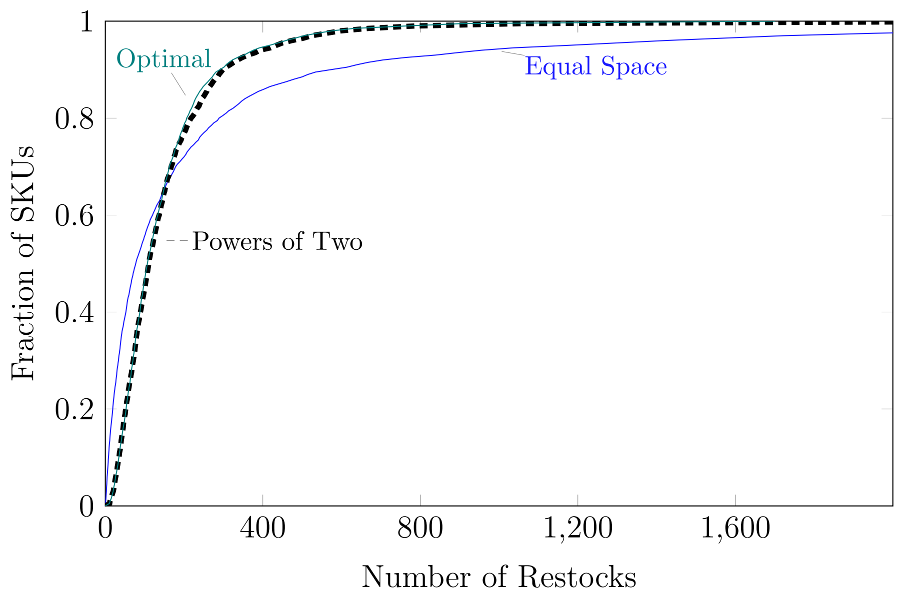

Figure 7: Cumulative distribution function of number of restocks of Equal Space, Optimal and Powers of Two schemes. The distribution for Optimal and Powers of Two tapers off much quicker than for Equal Space

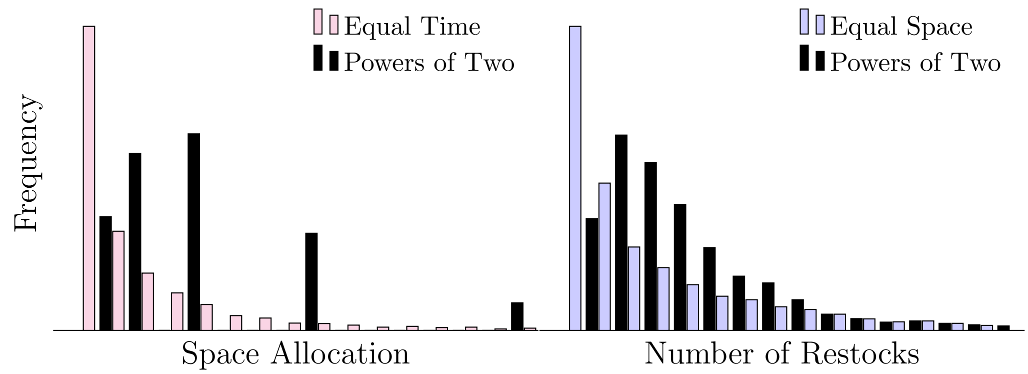

Figure 8: Histograms comparing space allocations of Powers of Two with Equal Time (left) and number of restocks required by Powers of Two scheme with Equal Space (right). In each case, the variability for Powers of Two is lower than the alternative.

However, not all SKUs are restocked at the same time like in the Equal Time allocation. But, being powers of two, the restocking intervals are synchronized allowing the warehouse manager to plan restocking activities. For example, if the minimum period is 7 days, and every SKU is either 7 days or 7 times a power of 2 days, then the warehouse manager can schedule workers who will restock the corresponding SKUs

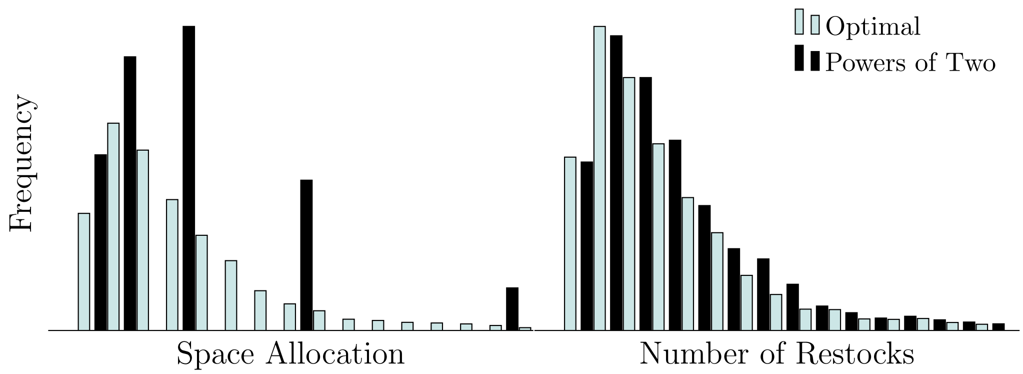

Figure 9: Histograms comparing space allocations (left) and number of restocks (right) of Powers of Two scheme with the Optimal allocation. The variability of the two are very similar with the added advantage of far fewer distinct values for space allocation.

on, say, every Friday. However, the space allocations generated need not be powers of two, requiring the shelves be organized to accommodate the sizes recommended by the algorithm.

The restocking time intervals are related to the number of restocks as follows. Let T be the time horizon, and t_i be the expected time between restocks for SKU i. The number of restocks for SKU i is f_i/v_i. We see that,

$$t_i = \frac{T}{f_i/v_i}$$

Therefore, if the number of restocks is expressed as a constant times a power of 2, immediately, the restocking time intervals are also a constant times a power of 2, since T, the time horizon is same for all SKUs. Therefore, we extend Algorithm 1 to determine an allocation such that the number of restocks for each SKU is represented as a constant factor times a power of two. We shall call this a space allocation with powers of two restocks. As before, this allocation produces no more than 1.06 times the optimal number of restocks.

As in Section 2.4, we define $z_i \in [1, 2)$ and an integer r_i for $i \in N$ such that,

$$\frac{f_i}{\tilde{v}_i} = z_i 2^{r_i}, \qquad \text{s.t } 1 \leq z_i < 2 \tag{2.8.1}$$

We construct a solution of the form $\mathbf{v}^k = (v_i^k, i \in N)$ where

$$\frac{f_i}{v_i^k} = \alpha^k 2^{q_i^k} \tag{2.8.2}$$

$$q_i^k = \begin{cases} r_i - 1 & i \leq k \\ r_i & i > k \end{cases} \tag{2.8.3}$$

and α^k is such that it minimizes $Z(\mathbf{v}^k)$. Since the capacity constraint (2.3.1) is tight, we determine α^k to be

$$\alpha^k = \frac{1}{V_F} \sum_{i \in N} \frac{f_i}{2^{q_i^k}} \tag{2.8.4}$$

Note that, since $r_i \geq q_i^k$ for all i and k and also that $\sum_{i \in N} \tilde{v}_i = \sum_{i \in N} \frac{f_i}{z_i 2^{r_i}} = V$,

$$\alpha^k = \frac{1}{V_F} \sum_{i \in N} \frac{f_i}{2^{q_i^k}} \geq \frac{1}{V_F} \sum_{i \in N} \frac{f_i}{2^{r_i}} \geq \min_i z_i \geq 1 \tag{2.8.5}$$

Algorithm 2 is a simple adaptation of Algorithm 1 to compute a space allocation with powers of two restocks.

Algorithm 2: Algorithm to round off the optimal number of restocks by SKU to powers of two.

---

1. Determine z_i and r_i for $i \in N$ using (2.8.1)
2. Index the SKUs as $i \in N = \{1, \ldots, n\}$, where $n = |N|$, such that $z_i \leq z_{i+1}$ for $1 \leq i \leq n-1$ and $n = |N|$
3. **for** $k = 1$ **to** n **do**
4. Determine $\mathbf{v}^k$ using (2.8.2), (2.8.3) and (2.8.4) to compute $Z(\mathbf{v}^k)$
5. Select the vector $\mathbf{v}^{\bar{k}}$ that has the least $Z(\mathbf{v}^{\bar{k}})$, $\bar{k} \in 1, \ldots, n$

Algorithm 2 runs in $O(|N| \log |N|)$ time since the work required in the algorithm is sorting in in Step 2 that takes $O(|N| \log |N|)$, and two $O(|N|)$ loops in Steps 1 and 3 respectively.

We show that this allocation has the exact same properties as the powers of two space allocations presented in Section 2.4. To do that, we first show the restock cost in both cases have the same functional form.

Lemma 2.8.1. The allocation with powers of two restocks produces a restock cost that is functionally identical to the powers of two space allocation.

Proof. As in Section 2.4, let

$$z_i^k = \begin{cases} z_i & i > k \\ 2z_i & i \leq k \end{cases} \tag{2.8.6}$$

Using definitions (2.8.6) and (2.8.3) in (2.8.1), we get

$$\tilde{v}_i = \frac{f_i}{z_i^k 2^{q_i^k}} \tag{2.8.7}$$

and therefore, writing v_i^k in terms of $\tilde{v}_i$ and z_i^k, using (2.8.2), (2.8.3), (2.8.4) and (2.8.7),

$$\begin{aligned} v_i^k &= \frac{f_i}{\alpha^k 2^{q_i^k}} && (2.8.8) \\ &= \frac{V_F}{\sum_{i \in N} f_i / 2^{q_i^k}} \tilde{v}_i z_i^k && (2.8.9) \\ &= \left(\frac{V_F}{\sum_{j=1}^n \tilde{v}_i z_i^k} \right) \tilde{v}_i z_i^k && (2.8.10) \end{aligned}$$

And hence rewriting the objective using (2.8.7), (2.4.12) and (2.4.13)

$$\begin{aligned} Z(\mathbf{v}^k) &= \sum_{i=1}^n \frac{f_i}{\tilde{v}_i} && (2.8.11) \\ &= \sum_{i=1}^n \frac{\sum_{j=1}^n \tilde{v}_j z_j^k}{V_F} \times \frac{f_i}{\tilde{v}_i z_i^k} && (2.8.12) \\ &= \sum_{i=1}^n \frac{\sum_{j=1}^n \tilde{v}_j z_j^k}{V_F} \frac{\lambda \tilde{v}_i}{z_i^k} && (2.8.13) \\ &= Z(\tilde{\mathbf{v}}) \sum_{i=1}^n \frac{\sum_{j=1}^n \tilde{v}_j z_j^k}{V_F^2} \frac{\tilde{v}_i}{z_i^k} && (2.8.14) \\ &= \frac{Z(\tilde{\mathbf{v}})}{V_F^2} \left\{ \sum_{i=1}^n \tilde{v}_i^2 + \sum_{i<j} \tilde{v}_i \tilde{v}_j \left(\frac{z_i^k}{z_j^k} + \frac{z_j^k}{z_i^k} \right) \right\} && (2.8.15) \end{aligned}$$

The rewritten objective, 2.8.15 is the same functional form as 2.4.14. □

Therefore, the solution has the same bound as the one for the powers of two space allocation.

Theorem 2.8.2. The allocation with powers of two restocks derived from Algorithm 2 costs at most 6% more than the optimal solution.

Similarly another important result, Theorem 2.5.4 that shows that none of the n candidate solutions generated cost more than 12.5% additional restocks over the optimal, also holds for the allocation with powers of two restocks. This is so because, the objective has the same functional form shown in (2.8.15) as the one in powers of two space allocation.

The results for allocation with powers of two restocking intervals and powers of two space allocations has an interesting parallel with the Equal Space and Equal Time allocations. As shown in Bartholdi and Hackman [7], the Equal Space and Equal Time allocations generate the same objective function and a very similar pattern is observed here. This allows the warehouse manager the flexibility to choose the type of allocation that suits the forward pick area (i.e. powers of two space allocation) or the warehouse work loads (i.e powers of two restocking intervals) but still not lose much over and above the optimal allocation.

2.9 Conclusions

The optimal allocations generated by the Hackman and Rosenblatt [23] lead to arbitrary divisions of space in the forward pick area. We show how to transform this space allocation into a more practical powers of two form and at the same time keep the number of restocks to no more than 6% over the optimal. The powers of two allocation is in the form of a constant times an integral power of two and therefore all allocations differ only by a power of two. The methodology allows us to use the ranking framework, where we can assign the SKUs based on its rank by labor efficiency to the forward pick areas and then determine a space allocation that is based on the powers of two methodology.

The algorithm produces as many allocations as the number of SKUs, each with a different constant. The best solution among those produced incurs no more than an additional 6% in restocking costs, and the worst incurs no more than an additional 12.5%. Therefore this allows some flexibility in choice of the constant term in the powers of two allocation depending on SKU and shelf geometries. The allocation also has very distinct values of space allocation for SKUs because the allocations belong to a finite set of multiples of a common base. In addition, we show that the variance of the powers of two space allocation and number of restocks is close to the variance of the optimal space allocation and restocks. These factors make the allocation *manageable* from the warehouse perspective without losing much compared to the optimal. Other commonly used allocation strategies like Equal Space and Equal Time not only have a much higher restocking cost but also have much higher variances than the optimal in the case of number of restocks (Equal Space) and space allocation (Equal Time).

We also see that when the number of SKUs is large, the choice of the constant becomes even more flexible, and one can choose *almost* any value between 1 and 2. Though choosing an allocation that is not the best among the ones produced by the algorithm increases the number of restocks, we know that that the increase is no more than 12.5% over the optimal.

Similarly, we also show how to allocate space to obtain powers of two restocking intervals. This produces a different form of manageability with restocking work being predictable and at the same time, not incurring any more than 6% restocking cost over and above the optimal. This is the counterpart of Equal Time and has the advantage of synchronizing restocking activities. Similar to the powers of two space allocation, results on variability, bounds on cost of the solutions and choice of values for the constant term hold good for the powers of two restocking intervals as well.

For a warehouse with over 3000 SKUs, the powers of two allocations generated by the algorithm are very close to optimal in key metrics. The cost of the allocations is no more than 4.2% from optimal, the variances of space allocations are no more than 17% than optimal and the variance of number of restocks is within 23% of the corresponding value for the optimal allocation. The corresponding metrics for Equal Space and Equal Time allocation strategies are significantly higher. The algorithm produces as many powers to two allocations as there are SKUs and the corresponding constant values so close together and fully spanning the interval $[1, 2)$. Therefore the warehouse can choose a desired value based on shelf geometries such that all allocations are multiples of the value and there exists a powers of two allocation with a constant term that is close the chosen value.

The powers of two allocation provides warehouse managers a choice of allocations to accommodate SKUs in forward pick areas using their chosen bin and shelf sizes. It is entirely possible that warehouse managers who prefer the advantages of Equal Space can potentially use powers of two allocations, and the managers who prefer Equal Time may benefit from the allocation with powers of two restocking intervals. Given that Equal Space and Equal Time are the predominant strategies employed by warehouses [8], the powers of two allocation can be a good choice for warehouses who seek to retain advantages of their present strategy and at the same time bring down their restocking costs closer to optimal.

Given a SKU assignment to a forward pick area, this chapter shows how to allocate space in a practical and manageable way but at the same time not very far from the optimal number of restocks determined by the fluid model. In the next three chapters we analyze the problem of determining the optimal assignment of SKUs to forward pick areas within a warehouse.

CHAPTER III

ASSIGNING SKUS IN A WAREHOUSE WITH ONE FORWARD PICK AREA

3.1 Introduction

Hackman and Rosenblatt [23] were the first to describe a mathematical model for the assignment and allocation of space to SKUs in a warehouse with one forward pick area. They employ a *fluid model* described in Chapter 1 to minimize the costs of picking and restocking, where the volume occupied by each SKU is continuously divisible and incompressible. They propose a heuristic to assign and allocate SKUs to the forward pick area that ranks SKUs on the basis of a ratio, labor efficiency, that can be computed a priori. The heuristic evaluates solutions such that SKUs assigned to the forward pick area are ranked higher than those that are not and chooses the one that minimizes total picking and restocking costs. Frazelle et al. [14] and later, Bartholdi and Hackman [7], show that the heuristic is near-optimal with the insight that higher ranked SKUs have a greater claim to the forward pick area. In this chapter, using their model shown in (1.3.1), we show an alternative proof of the same results using the primal formulation. Given any SKU assignment that does not adhere to the ranking, i.e SKUs assigned to the forward pick area ranked lower than one or more SKUs that are not, we show how to switch a pair of SKUs where a higher ranked SKU replaces the lower ranked one assigned to the forward pick area. The switching of SKUs leads to an assignment and allocation solution that improves the total benefit while maintaining feasibility. This approach enables us to extend the same results in several ways shown in subsequent chapters.

3.2 *The mathematical model*

We consider the case of a warehouse with one forward pick area. Using the same notation and representation as in Chapter 1, we restate the formulation of the fluid model in (1.3.1 to assign and allocate SKUs to the forward pick area.

$$\text{OPT}: \max_{\mathbf{x},\mathbf{v}} \sum_{i \in N} x_i \left(sp_i - \frac{c_r f_i}{v_i} \right) \tag{3.2.1a}$$

subject to

$$\sum_{i \in N} x_i v_i \leq V_F \tag{3.2.1b}$$

$$x_i \in \{0, 1\} \qquad i \in N, \tag{3.2.1c}$$

$$v_i \geq 0 \qquad i \in N \tag{3.2.1d}$$

From Hackman and Rosenblatt [23] and expression (1.2.2), we know that if given the assignment, $\mathbf{x}$, the optimal allocation of space to SKU i is proportional to the square root of its flow, f_i as shown in (3.2.2):

$$\tilde{v}_i(\mathbf{x}) = \frac{\sqrt{f_i}}{\sum_{j \in N} x_j \sqrt{f_j}} V_F. \tag{3.2.2}$$

Following Frazelle [14], we eliminate the $\mathbf{v}$ variable in the formulation by using the optimal value for v_i in OPT to get

$$\max_{\mathbf{x}} Z(\mathbf{x}) \equiv \sum_{i \in N} x_i sp_i - \frac{c_r}{V_F} \left(\sum_{i \in N} x_i \sqrt{f_i} \right)^2 \tag{3.2.3a}$$

subject to

$$x_i \in \{0, 1\} \qquad i \in N \tag{3.2.3b}$$

This problem as formulated is NP-Hard. In order to prove that it is, we reduce the subset sum problem, which is NP-Complete, to the problem as defined in (3.2.3). The subset sum problem is defined as follows: Given a set of integers $S = \{a_1, a_2, \ldots, a_n\}$, and an integer t, determine a subset $S' \subseteq S$, if one exists, such that

$$\sum_{i: a_i \in S'} a_i = t.$$

Formulating this as an integer program, we get:

$$\min_{\mathbf{x}} \left(t - \sum_{1 \leq i \leq n} x_i a_i \right)^2 \tag{3.2.4a}$$

subject to

$$x_i \in \{0, 1\} \qquad\qquad 1 \leq i \leq n \tag{3.2.4b}$$

If the solution, $\tilde{\mathbf{x}}$, to the problem results in the objective (3.2.4a) evaluating to zero, then there is a subset $S' = a_{i\tilde{x}_i=1}$ that sums to t. Now, rewriting the objective function (3.2.4a) by expanding the squares and converting to a maximization problem, we get:

$$-t^2 + \max_{\mathbf{x}} 2t \sum_{1 \leq i \leq n} x_i a_i - \left(\sum_{1 \leq i \leq n} x_i a_i \right)^2 \tag{3.2.5a}$$

subject to

$$x_i \in \{0, 1\} \qquad\qquad 1 \leq i \leq n \tag{3.2.5b}$$

We make the following assignment that reduces (3.2.5) to an instance of (3.2.3):

$$s = 2t$$

$$p_i = a_i$$

$$\sqrt{f_i} = a_i.$$

If there is a polynomial algorithm that solves (3.2.3), then the subset problem can also be solved using the same algorithm contradicting the fact that it is NP-Complete. Thus, the problem of assignment and allocation of SKUs in the forward pick area is NP-Hard.

3.3 A proof based on switching SKUs

Hackman and Rosenblatt [23] propose a ranking-based heuristic to the formulation to determine an assignment of SKUs to the forward pick area such that they are ranked higher than those that are not. This requires the evaluation of $|N|$ such solutions that satisfies the ranking criteria and choosing the best among them. This heuristic runs in polynomial time and Frazelle et al. [14] show that solution produced is within the net benefit of one SKU from optimal. Typical warehouses stock hundreds, if not thousands of skus, making the solution very close to optimal. In this section, we present an alternate proof for the result via the primal formulation that is based on a switching argument where given any solution that does not follow the ranking criteria, it can be improved by swapping in the higher ranked unassigned SKU with the lower ranked assigned one. In addition to providing an insight into how to improve any given solution, this is also extensible to the generalizations of the models we consider in subsequent chapters.

Following Bartholdi and Hackman [7], we define the *labor efficiency* of a SKU i to be

$$\frac{p_i}{\sqrt{f_i}}.$$

We rank the SKUs by the descending order of their labor efficiencies, breaking ties arbitrarily.

We relax the binary constraint on $\mathbf{x}$ and show for the relaxed problem $\mathrm{OPT}-R$ shown in (3.3.1) below that the SKUs that go into forward pick area F are the ones with the greatest labor efficiency.

$$\mathrm{OPT}-R : \max_{\mathbf{x}} Z(\mathbf{x}) \equiv \sum_{i\in N} x_i s p_i - \frac{c_r}{V_F}\left(\sum_{i\in N} x_i\sqrt{f_i}\right)^2 \tag{3.3.1a}$$

subject to

$$0 \leq x_i \leq 1 \qquad\qquad i \in N \tag{3.3.1b}$$

Theorem 3.3.1. The SKUs that have a claim to forward pick area F are those with the highest rank. That is, there exists an optimal solution, $\tilde{\mathbf{x}}$, for the relaxed problem $\mathrm{OPT}-R$ where, if SKU i outranks SKU j then $\tilde{x}_i \geq \tilde{x}_j$ and more specifically, if $\tilde{x}_j > 0$ then $\tilde{x}_i = 1$.

Proof. Consider an optimal solution $\tilde{\mathbf{x}}$ to the relaxed problem OPT$-R$. If the solution satisfies the condition in the theorem, we are done. If not, then there exists two SKUs i and j where, i outranks j and $\tilde{x}_j > \tilde{x}_i$. We modify this solution to construct an alternate optimal solution where SKU i is switched with SKU j in the forward pick area and gives us the desired properties.

We consider a perturbation of the vector $\tilde{\mathbf{x}}$ by an $\epsilon > 0$ to obtain a new assignment vector $\bar{\mathbf{x}}$ in the following manner:

$$\bar{x}_k = \begin{cases} \tilde{x}_k & k \in N \setminus \{i,j\} \\ \tilde{x}_i + \epsilon\sqrt{f_j}/\sqrt{f_i} & k = i \\ \tilde{x}_j - \epsilon & k = j \end{cases}$$

We suppress the functional dependence of $\bar{\mathbf{x}}$ on ϵ. Since by assumption $\tilde{x}_j > \tilde{x}_i$, a positive choice of ϵ exists that ensures feasibility. This particular switching ensures that for any feasible choice of ϵ the total restocking costs associated with the new assignment vector $\bar{\mathbf{x}}$ is identical to the total restocking costs associated with the given optimal assignment vector $\tilde{\mathbf{x}}$ as verified by the following equalities:

$$\begin{aligned} \frac{c_r}{V_F}\left(\sum_{k\in N}\bar{x}_k\sqrt{f_k}\right)^2 &= \frac{c_r}{V_F}\left(\sum_{k\neq i,j}\bar{x}_k\sqrt{f_k} + \bar{x}_i\sqrt{f_i} + \sum_{k\in N}\bar{x}_j\sqrt{f_j}\right)^2 \\ &= \frac{c_r}{V_F}\left[\sum_{k\neq i,j}\tilde{x}_k\sqrt{f_k} + \left(\tilde{x}_i + \epsilon\frac{\sqrt{f_j}}{\sqrt{f_i}}\right)\sqrt{f_i} + \sum_{k\in N}(\tilde{x}_j + \epsilon)\sqrt{f_j}\right]^2 \\ &= \frac{c_r}{V_F}\left(\sum_{k\in N}\tilde{x}_k\sqrt{f_k}\right)^2. \end{aligned}$$

Moreover, the pick savings associated with $\bar{\mathbf{x}}$,

$$\begin{aligned} \sum_{k\in N}\bar{x}_k s p_k &= \sum_{k\neq i,j}\tilde{x}_k s p_k + s p_i\left(\tilde{x}_i + \epsilon\frac{\sqrt{f_j}}{\sqrt{f_i}}\right) + s p_j(\tilde{x}_j - \epsilon) \\ &= \sum_k \tilde{x}_k s p_k + \epsilon s\left(p_i\frac{\sqrt{f_j}}{\sqrt{f_i}} - p_j\right), \end{aligned}$$

is at least as large as the pick savings associated with $\tilde{\mathbf{x}}$, since $\frac{p_i}{\sqrt{f_i}} \geq \frac{p_j}{\sqrt{f_j}}$. Since $\tilde{\mathbf{x}}$ is an optimal solution and restocking costs are equal, it must be the picking costs are equal. Therefore $\bar{\mathbf{x}}$ is also an optimal solution to OPT $- R$. This demonstrates that it is always possible to construct a new optimal solution from an existing optimal solution that has the requisite ranking property, which establishes the first part of the theorem. The second part of the theorem is established by choosing the *maximum* possible ϵ that ensures feasibility, namely,

$$\max\left\{\epsilon : \tilde{\mathbf{x}} + \epsilon\frac{\sqrt{f_j}}{\sqrt{f_i}} \leq 1, \tilde{x}_j - \epsilon \geq 0\right\}.$$

The maximal ϵ is such that the values $(\bar{x}_i, \bar{x}_j)$ lie on the boundary of the unit square; in particular, $(\bar{x}_i, \bar{x}_j)$ will either be equal to $(1, \bar{x}_j)$ or $(\bar{x}_i, 0)$ or $(1, 0)$, which proves the second part that if $\bar{x}_j > 0$ then $\bar{x}_i = 1$. $\square$

Corollary 3.3.2 immediately follows from Theorem 3.3.1.

Corollary 3.3.2. There exists an optimal solution, $\tilde{\mathbf{x}}$, to problem (3.2.3) for which there is at most one fractional SKU, that is a SKU with corresponding assignment $x \in (0, 1)$. It has the least rank among those with $x > 0$. Any SKU i with $\tilde{x}_i = 0$ is of a lesser rank than the fractional SKU.

Proof. From Theorem 3.3.1. $\square$

Theorem 3.3.1 and Corollary 3.3.2 prove that there is a solution such that for any SKU i such that $\tilde{x}_i > 0$ outranks all SKUs j with $\tilde{x}_j = 0$. Moreover there is at most only 1 SKU i such that $\tilde{x}_i \in (0, 1)$, i.e. fractional, and that all SKUs j with $\tilde{x}_j = 1$ outrank any such fractional SKU. From such an optimal solution $\tilde{\mathbf{x}}$, we construct a solution to the original problem, $\mathbf{x}^*$ such that,

$$x_i^* = \begin{cases} 1 & \tilde{x}_i = 1 \\ 0 & \text{otherwise} \end{cases}$$

Theorem 3.3.3. The solution $\mathbf{x}^*$ is within the profit of one SKU from the optimal value of OPT.

Proof. Let $\bar{\mathbf{x}}$ be the optimal assignment vector to problem OPT. As defined earlier, $\tilde{\mathbf{x}}$ is the optimal solution to OPT $- R$. Let $k \in N$ be the SKU for which $\tilde{\mathbf{x}}$ is fractional. Since $\mathbf{x}^*$ is a feasible solution to problem OPT and OPT $- R$ is a relaxation of OPT, we get

$$Z(\mathbf{x}^*) \leq Z(\bar{\mathbf{x}}) \leq Z(\tilde{\mathbf{x}}).$$

Hence,

$$\begin{aligned}
Z(\bar{\mathbf{x}}) - Z(\mathbf{x}^*) &\leq Z(\tilde{\mathbf{x}}) - Z(\mathbf{x}^*) \\
&\leq x_k^* sp_k - \frac{c_r}{V_F}\left[x_k^{*2} f_k + 2x_k^* \sqrt{f_k}\left(\sum_{i:x_i^*=1} \sqrt{f_i}\right)\right] && (3.3.2) \\
&\leq x_k^*\left[sp_k - \frac{c_r f_k}{V_F\sqrt{f_k}}\left(x_k^*\sqrt{f_k} + \sum_{i:x_i^*=1} \sqrt{f_i}\right)\right] && (3.3.3) \\
&\leq x_k^*\left(sp_k - \frac{c_r f_k}{\tilde{v}_k(\tilde{\mathbf{x}})}\right), && (3.3.4)
\end{aligned}$$

which is the profit of one SKU. □

In warehouses with significant each picking, the number of SKUs, $|N|$, is generally large. Therefore, in such cases, the bound on the total net benefit produced by the assignment $\mathbf{x}^*$ being within the net benefit of one SKU from optimal is a very tight one.

To obtain the solution $\mathbf{x}^*$, we present Algorithm 3 that was originally shown in Hackman and Rosenblatt [23].

Corollary 3.3.4. Algorithm 3 requires a computational effort of at most $|N| \log |N| + |N|^2$

Proof. The time taken to sort the $|N|$ SKUs is $|N| \log |N|$ and the loop in step 5 of Algorithm 3 runs $|N|$ times, with at most $|N|$ calculations within each loop. □

The algorithm for determining a near-optimal solution for the case of one forward mode is computationally efficient. The performance of the algorithm can be improved

Algorithm 3: Algorithm to determine assignment and allocation of SKUs for a warehouse with one forward pick area.

Sort SKUs in set N in the descending order of their labor efficiencies, breaking ties arbitrarily.
Set $S \leftarrow 0$.
for $k = 1$ **to** $|N|$ **do**
    Assign the first k SKUs to the forward pick area F by setting $x_i \leftarrow 1$ for those k SKUs and $x_i \leftarrow 0$ for the rest.
    Determine optimal allocation, $\mathbf{v}$, for assignment $\mathbf{x}$ using expression (3.2.2).
    For such an assignment, determine the objective value, $Z(\mathbf{x})$, from the expression (3.2.3a).
    Set $S \leftarrow \max\{S, Z(\mathbf{x})\}$
Return S and the assignment that produces it. If $S = 0$ then the assignment is the zero $\mathbf{x}$ vector.

further to $O(|N| \log |N|)$ time by only calculating the values of the additional SKU in expression (3.2.3a). At each iteration of the loop, we maintain two variables, the first is the sum of picks and the second, the sum of square root of flow of SKUs selected so far. These two variables are sufficient to determine the total net benefit of the SKUs selected so far to be assigned to the forward pick area. This takes only constant time, and therefore the loop time is only $O(|N|)$. Consolidating with time required for the initial sorting of SKUs in descending order of labor efficiency, we get a total time of $O(|N| \log |N|)$. Therefore, for not much more than the time to sort the SKUs, we can determine a near-optimal solution to the assignment and allocation problem of SKUs.

In addition to Theorem 3.3.1 being useful to show the result of Hackman and Rosenblatt, we also gain a valuable insight on how to improve any solution that may deviate from the ranking suggested by the Hackman and Rosenblatt [23] heuristic. Given minor changes in SKUs or demand patterns, we see how to replace a lower ranked SKU assigned to the forward pick area with a higher ranked one as long as the square root of the flow of the leaving SKU can be substituted by the entering SKU. An equivalent switching argument can be made also by modifying the solution to keep picking costs the same to get an equivalent or lower restock cost that maybe easier to understand and use.

3.4 Conclusions

The fluid model and the solution procedure of Hackman and Rosenblatt [23] is a very effective way to solve the assignment and allocation problem. The heuristic in Algorithm 3 is computationally efficient and simple to implement. More importantly, the ranking of SKUs by labor efficiency that decides which SKUs are assigned to the forward pick area is a critical insight that allows warehouse managers to discern immediately by using a single metric as to which SKUs have a higher claim to the forward pick area.

Frazelle et al. [14] and later, Bartholdi and Hackman [7] showed that the Hackman and Rosenblatt [23] heuristic produces a near-optimal solution. In this chapter, we employ a new argument based on switching SKUs assigned to the forward pick area to re-create the same result. This has several advantages. In proving the result, we gain a useful insight on how to improve a solution that deviates from the ranking-based assignment to the forward pick area by switching SKUs. Additionally, the methodology can be easily adapted to show a similar result when the cost parameters, cost per pick and cost per restock, vary by SKU. And finally, we use the approach in subsequent chapters to extend and generalize the fluid model and derive a ranking-based algorithm to construct a near-optimal solutions in each of those cases.

CHAPTER IV

ASSIGNING SKUS IN A WAREHOUSE WITH MULTIPLE FORWARD PICK AREAS AND CONSTRAINTS

4.1 Introduction

In Chapter 3, we consider a warehouse with one forward pick area and review the Hackman and Rosenblatt [23] algorithm to assign SKUs and allocate space. Using a new argument, we re-create the result shown by Frazelle et al. [14] that the algorithm produces a solution that is near-optimal via the primal formulation. In Chapter 2 we show that when allocating space in forward pick areas we can accommodate the geometry of SKUs and storage units and still be very close to the optimal as determined by the fluid model.

A natural extension to consider is a warehouse with multiple forward pick areas. This occurs frequently in warehouses, where the warehouse may locate disjoint forward pick areas in multiple locations, thereby each having its own space constraints and cost structure on picks and restocks made. Or, the warehouse may use multiple types of storage units like flow racks, shelving racks, Automated Storage and Retrieval System (AS/RS) among others and each type of storage unit differs in the space available to store SKUs and the costs incurred by picking and restocking activities. To assign and allocate space in multiple forward pick areas, Bartholdi and Hackman [6] extend the ranking methodology and derive an algorithm that generates a near-optimal solution using the fluid model. We extend the new argument presented in Chapter 3 to derive a similar result for the warehouse with multiple forward pick areas.

Warehouses, in practice, also have limitations on labor available for picking and restocking activities. Frazelle et al. [14] extend the Hackman and Rosenblatt [23] fluid model to include a constraint on congestion within the forward pick area that is expressed as a function of the total time spent picking and restocking the forward pick area. They show how to extend the ranking-based algorithm to determine a near-optimal assignment and allocation of SKUs for a warehouse with one forward pick area. In addition to the constraint on congestion, we consider limits on number of restocks that can be performed and the number of picks that can be made in a time period because of layout of the warehouse, limited labor availability or equipment used to pick and restock. For example, a warehouse that shares workers or aisles for picking and restocking, may plan on using a portion of the day for restocks and the rest for picking. This will mean that only a certain number of picks and restocks can be accomplished in any given time period. Equipment such as carousels and conveyors also place limits on picking and restocking activities. We show how to extend the ranking framework to determine a near-optimal assignment and allocation for warehouses with such constraints on one forward pick area. We also show how to extend the same algorithm for warehouses with multiple forward pick areas. To derive the results in each case, we use an argument similar to the single forward pick area case presented in Chapter 3.

4.2 Extending the fluid model

We consider the case of a warehouse with multiple forward pick areas use a cost structure for picks and restocks similar to the case of the warehouse with one forward pick area in Chapter 3. Each forward pick area has a constant pick savings for each pick relative to the reserve and incurs a fixed restocking cost to replenish SKUs from the reserve.

An example with two forward pick areas, F_1 and F_2 are illustrated in Figure 10.

Each forward pick area has a space constraint and each pick from the forward pick area has a comparative cost savings over the reserve represented by s_1 and s_2. The forward pick areas have a cost to restock and each trip from the reserve incurs a cost of c_{r1} and c_{r2} for F_1 and F_2 respectively. The question we address in this section is whether a SKU is to be picked from a forward pick area, and if so, which one and how much space should be allocated to the SKU in that forward pick area.

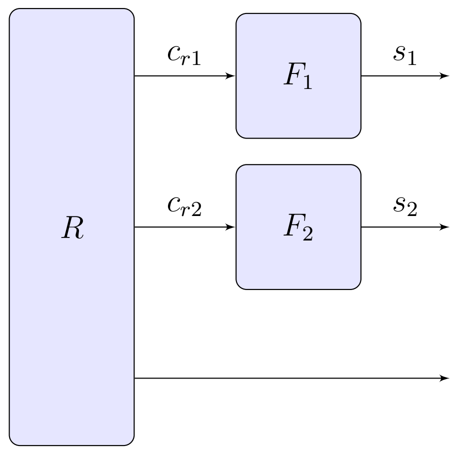

Figure 10: A warehouse with two forward pick areas, F_1 and F_2. The savings in pick cost for F_1 and F_2 relative to the reserve are s_1 and s_2 respectively. The corresponding restocking costs for each forward pick area are c_{r1} and c_{r2}.

To construct the fluid model, we extend the notation defined in Chapter 3 for a warehouse with a single forward pick area to one with multiple forward pick areas. Let M be the set of forward pick areas and for $m \in M$, let s_m be the savings in picking from forward pick area m relative to picking from the reserve. Let c_{rm} be the restocking cost incurred each time any SKU has to be restocked to forward pick area m irrespective of the quantity. Let V_m be the volume of forward pick area m.

The decision variables for the multiple forward pick areas as follows. We use a binary decision variable, x_{mi}, $m \in M$, $i \in N$, that is 1 if SKU i is assigned to forward pick area m and 0 otherwise. The continuous decision variable, v_{mi} is the volume of space allocated to SKU i in forward pick area m. We formulate the optimization problem for multiple forward pick areas, OPT_M, as shown in (4.2.1).

$$\text{OPT}_M : \max_{\mathbf{x},\mathbf{v}} \sum_{m\in M} \sum_{i\in N} x_{mi} \left(s_m p_i - \frac{c_{rm} f_i}{v_{mi}} \right) \tag{4.2.1a}$$

subject to

$$\sum_{i\in N} x_{mi} v_{mi} \leq V_m \qquad \forall m \in M \tag{4.2.1b}$$

$$\sum_{m\in M} x_{mi} \leq 1 \qquad \forall i \in N, \tag{4.2.1c}$$

$$x_{mi} \in \{0,1\} \qquad m \in M, i \in N, \tag{4.2.1d}$$

$$v_{mi} \geq 0 \qquad m \in M, i \in N \tag{4.2.1e}$$

Notice that when given the assignment $\mathbf{x}$ of SKUs to the individual forward pick areas, the allocation problem reduces to the single forward pick area problem for each forward pick area. Hence, the allocation, as in the single forward pick area problem, is proportional to the square root of the flow. That is,

$$v_{mi} = \frac{\sqrt{f_i}}{\sum_{j\in N} x_{mj} \sqrt{f_j}} V_m. \tag{4.2.2}$$

Reformulating the problem by eliminating v_{mi} in OPT_M shown in (4.2.1) we get an equivalent formulation,

$$\text{OPT}_M : \max_{\mathbf{x}} Z(\mathbf{x}) \equiv \sum_{m\in M} \left[\sum_{i\in N} x_{mi} s_m p_i - \frac{c_{rm}}{V_m} \left(\sum_{i\in N} x_{mi} \sqrt{f_i} \right)^2 \right] \tag{4.2.3a}$$

subject to

$$\sum_{m\in M} x_{mi} \leq 1 \qquad \forall i \in N, \tag{4.2.3b}$$

$$x_{mi} \in \{0,1\} \qquad m \in M, i \in N \tag{4.2.3c}$$

This formulation is a generalized resource allocation problem, where the forward pick areas are resources with capacities to which SKUs are assigned. Each SKU can be assigned to only one forward pick area and the objective function is additively separable by SKU. Given that there are typically many SKUs in a warehouse with each SKU allocated a space much smaller than total capacity of any forward pick area, this problem lends itself to be solved by the non-smooth convex optimization approach proposed by Hackman and Platzman [22]. Their algorithm under the given problem conditions, produces a solution that is near-optimal in polynomial time. However, the application of this general approach has two disadvantages. First, their solution procedure requires a sophisticated algorithm that iterates over non-smooth convex relaxations of the problem. Second, more importantly, is that we lose the ranking framework that is present in the single forward pick area case. In the next section, we extend the ranking-based methodology developed in Chapter 3 to derive an algorithm that solves the problem of assigning and allocating SKUs to a warehouse with multiple forward pick areas.

4.3 Assignment and allocation of SKUs to multiple forward pick areas

Like in the case of the single forward pick area, we rank the SKUs in descending order of labor efficiency. In addition to ranking the SKUs, we also rank the forward pick areas in the descending order of pick savings. As before, we break ties arbitrarily. The combination of the rankings for SKUs and forward pick areas produces the necessary insight to manage the assignment of SKUs to multiple forward pick areas. Bartholdi and Hackman [6] show that high rank SKUs have a greater claim to the high rank forward pick areas than do lower rank SKUs. In this section, we show the same result by analyzing the primal formulation and extending the proof technique presented in Chapter 3.

We first show the ranking property for the linear relaxation of the problem. By relaxing the binary constraint on $\mathbf{x}$, we obtain the relaxed problem $\mathrm{OPT}_M - R$ shown in (4.3.1).

$$\text{OPT}_M - R : \max_{\mathbf{x}} Z(\mathbf{x}) \equiv \sum_{m \in M} \left[\sum_{i \in N} x_{mi} s_m p_i - \frac{c_{rm}}{V_m} \left(\sum_{i \in N} x_{mi} \sqrt{f_i} \right)^2 \right] \tag{4.3.1a}$$

subject to

$$\sum_{m \in M} x_{mi} \leq 1 \qquad \forall i \in N, \tag{4.3.1b}$$

$$0 \leq x_{mi} \leq 1 \qquad m \in M, i \in N \tag{4.3.1c}$$

Theorem 4.3.1. Let $m, w \in M$, be two forward pick areas such that m outranks w. Then there is an optimal solution, $\tilde{\mathbf{x}}$, where for two SKUs $i, j \in N$ such that $\tilde{x}_{mi} > 0$ and $\tilde{x}_{wj} > 0$, SKU i outranks SKU j. Further, in that solution, at least one of $\tilde{x}_{mj}$ and $\tilde{x}_{wi}$ are zero.

Proof. Suppose there does not exist such a solution. Let $\tilde{\mathbf{x}}$ be an optimal solution, where, for two forward pick areas $m, w \in M$, two SKUs $i, j \in N$, m outranks w, $\tilde{x}_{mj} > 0$ and $\tilde{x}_{wi} \geq 0$ and i outranks j. Modifying the optimal solution by switching SKUs, we show how to construct alternative solutions that satisfy the required properties and contradict the assumption.

Consider the case where $\tilde{x}_{wi} = 0$. This implies there is no forward pick area w where SKU i is stored, or equivalently, SKU i which is higher ranked than SKU j is assigned to the reserve. As in Theorem 3.3.1, construct an alternate feasible solution $\bar{\mathbf{x}}$ by modifying $\tilde{\mathbf{x}}$ as follows:

$$\bar{x}_{mk} = \begin{cases} \tilde{x}_{mk} & k \in N \setminus \{i, j\} \\ \tilde{x}_{mi} + \epsilon \sqrt{f_j} / \sqrt{f_i} & k = i \\ \tilde{x}_{mj} - \epsilon & k = j \end{cases}$$

As shown in Theorem 3.3.1, this solution is feasible and also has the same restocking cost as the optimal assignment $\tilde{\mathbf{x}}$. In addition, the pick savings is at least as much as the optimal as $p_i/\sqrt{f_i} \geq p_j/\sqrt{f_j}$. Therefore the total net benefit for the new assignment $\bar{\mathbf{x}}$ is

$$Z(\bar{\mathbf{x}}) = \sum_{l\in M, k\in N} \tilde{x}_{lk} s p_k + \epsilon s \left(p_i\sqrt{f_j}/\sqrt{f_i} - p_j\right) - \sum_{l\in M} \frac{c_r}{V_m} \left(\sum_{k\in N} \tilde{x}_{lk}\sqrt{f_k}\right)^2.$$

Following Theorem 3.3.1, we choose the maximum possible value for ϵ and therefore, either $(\bar{x}_{mi}, \bar{x}_{mj})$ is $(1, \bar{x}_{mj})$ or $(\bar{x}_{mi}, 0)$. This contradicts our assumption that no such solution exists. Therefore, it must be that $\tilde{x}_{wi} > 0$ for some forward pick area w that is outranked by forward pick area m.

Using the optimal assignment $\tilde{\mathbf{x}}$ such that $\tilde{x}_{wi} > 0$ for some forward pick area w, construct an alternate solution, $\bar{\mathbf{x}}$ as follows:

$$\bar{x}_{lk} = \begin{cases} \tilde{x}_{lk} & l \in M \setminus \{m, w\} \text{ and } k \in N \setminus \{i, j\} \\ \tilde{x}_{mj} - \epsilon & l = m, k = j \\ \tilde{x}_{mi} + \epsilon\sqrt{f_j}/\sqrt{f_i} & l = m, k = i \\ \tilde{x}_{wj} + \epsilon & l = w, k = j \\ \tilde{x}_{wi} - \epsilon\sqrt{f_j}/\sqrt{f_i} & l = w, k = i \end{cases}$$

Observe that there is a $\epsilon > 0$ for which the solution $\bar{\mathbf{x}}$ is feasible since $\bar{x}_{mj} > 0$ and $\bar{x}_{wi} > 0$. Choose the maximum possible ϵ such that the solution is feasible. Observe that at least one of $\bar{x}_{mj}$ or $\bar{x}_{wi}$ is zero. Hence the total net benefit $Z(\bar{\mathbf{x}})$ evaluates to,

$$\begin{aligned} Z(\bar{\mathbf{x}}) = &\sum_{l\in M} \left[\sum_{k\in N} \tilde{x}_{lk} s_l p_k - \frac{c_{rl}}{V_l} \left(\sum_{k\in N} \tilde{x}_{lk}\sqrt{f_k}\right)^2\right] \\ &+ \epsilon \left[s_m(p_i\sqrt{f_j}/\sqrt{f_i} - p_j) + s_w(-p_i\sqrt{f_j}/\sqrt{f_i} + p_j)\right]. \end{aligned}$$

Note that the new solution incurs the same restocking cost, $\left(\sum_{k\in N} \tilde{x}_{lk}\sqrt{f_k}\right)^2$. Furthermore,

$$Z(\bar{\mathbf{x}}) = Z(\tilde{\mathbf{x}}) + \frac{\epsilon}{\sqrt{f_j}}(s_m - s_w)(p_i/\sqrt{f_i} - p_j/\sqrt{f_j}).$$

Now the term, $(s_m - s_w)(p_i/\sqrt{f_i} - p_j/\sqrt{f_j})$ is non-negative since m outranks w and i outranks j. Hence,

$$Z(\bar{\mathbf{x}}) \geq Z(\tilde{\mathbf{x}}).$$

The solution $\bar{\mathbf{x}}$ satisfies the requirements of the theorem contradicting the assumption of non-existence of such a solution. $\square$

With this Theorem, we next show that the higher ranked forward pick areas claim the higher ranked SKUs.

Corollary 4.3.2. There exists an optimal solution to problem (4.3.1) such that the SKUs assigned to a higher ranked forward pick area outrank those assigned to a lower ranked forward pick area. Furthermore, in that solution the assignment variables, x, corresponding to at most $|M|$ SKUs are fractional, that is for such a SKU i have assignment $x_{mi} \in (0, 1)$ for some forward pick area m.

Proof. From Theorem 4.3.1, we see that there exists an optimal solution of the following form: If the SKUs are ordered in descending order of labor efficiency, and partitioned into $|M| + 1$ sets, allowing for fractional division of SKUs across partitions. The first partition is allocated to the highest ranked forward pick area, i.e, the one with the highest pick savings, s, the next partition to the next ranked forward pick area and so on with the last partition allocated to the reserve. Therefore, the first part of the corollary immediately follows that if a SKU is assigned to a higher ranked forward pick area than another SKU, then it must necessarily outrank that SKU.

For this solution, among the SKUs allocated to the highest ranked forward pick area, say, m, there can be only one fractional SKU: the least ranked SKU. The lowest rank SKU, i, in forward pick area m, if fractional, i.e., $\tilde{x}_{mi} < 1$, then, the highest ranked SKU in the next ranked forward pick area to m (if one exists), say n, is SKU i, with $\tilde{x}_{wi} \leq 1 - \tilde{x}_{mi}$. If SKUs ranked lower than i are assigned to any of the forward pick areas, it must be that $\tilde{x}_{wi} = 1 - \tilde{x}_{mi}$ since a higher ranked SKU must be assigned before a lesser one. For the forward pick areas ranked 2 to $|M|$, there can be only at most two SKUs that are fractionally allocated: the ones of highest and lowest rank among those allocated to the forward pick area. The highest rank SKU is shared with the next higher ranked forward pick area and so the total number of fractional SKUs is $|M|$. □

As before, we construct a solution to the original problem, $\mathbf{x}^*$ such that,

$$x_i^* = \begin{cases} 1 & \tilde{x}_i = 1 \\ 0 & \text{otherwise} \end{cases}$$

This solution is the near-optimal solution we seek.

Corollary 4.3.3. The solution, $\mathbf{x}^*$, is at most profit of M SKUs from optimal.

Proof. Let $\bar{\mathbf{x}}$ be the optimal assignment vector to problem OPT_M. As defined earlier, $\tilde{\mathbf{x}}$ is the optimal solution to the relaxed problem $\mathrm{OPT}_M - R$. Let $K \subseteq N$ be the set of SKUs for which $\tilde{\mathbf{x}}$ is fractional. We know from Corollary 4.3.2, that $|K| \leq |M|$. For every SKU, $k \in K$, let w_k be the set of forward pick areas for which $\tilde{x}_{mk} > 0$, where $m \in w_k$. By construction of $\mathbf{x}^*$ we know that

$$Z(\mathbf{x}^*) \leq Z(\bar{\mathbf{x}}) \leq Z(\tilde{\mathbf{x}}).$$

Hence, the difference between the optimal and the constructed solution is

$$\begin{aligned} Z(\bar{\mathbf{x}}) - Z(\mathbf{x}^*) &\leq Z(\tilde{\mathbf{x}}) - Z(\mathbf{x}^*) \\ &\leq \sum_{k \in K, m \in w_k} x_{mk}^* s_m p_k - \frac{c_{rm}}{V_m} \left(x_{mk}^{*2} f_k + 2 x_k^* \sqrt{f_k} \sum_{i : x_{mi}^* > 0} \sqrt{f_i} \right) \\ &\leq \sum_{k \in K, m \in w_k} x_{mk}^* \left\{ s_m p_k - \frac{c_{rm} f_k}{V_m \sqrt{f_k}} \left(x_{mk}^* \sqrt{f_k} + 2 \sum_{i : x_{mi}^* = 1} \sqrt{f_i} \right) \right\} \\ &\leq \sum_{k \in K} \max_{m \in w_k} \left\{ s_m p_k - \frac{c_{rm} f_k}{V_m \sqrt{f_k}} \left(\sqrt{f_k} + 2 \sum_{i : x_{mi}^* = 1} \sqrt{f_i} \right) \right\}. \end{aligned}$$

Since $K \leq |M|$, we see that that the difference is within the profit of $|M|$ SKUs.

□

Usually, the number of SKUs in warehouses with each picking is very large and hence the bound of being within the profit of at most $|M|$ SKUs from optimal is a very tight bound. Algorithm 4 shows how to construct the near-optimal solution, $\mathbf{x}^*$.

Algorithm 4: Algorithm to determine SKU assignment and allocations for a warehouse with multiple forward pick areas.

Sort the SKUs in the set N in the descending order of their labor efficiencies, breaking ties arbitrarily

Sort the forward pick areas in the set M in the descending order of their pick savings, breaking ties arbitrarily. Let the forward pick areas be $M = \{1, 2, \dots, |M|\}$.

Set $S \leftarrow 0$

foreach $\mathbf{k} = (k_1, k_2, \dots, k_{|M|})$ *ordered* $|M|$*-tuple such that* $\sum_{m \in M} k_m \leq |N|$ *and* $k_m \in \{0, 1, \dots, |N|\}$ **do**

> Select the first k_1 SKUs and assign them to forward pick area 1, k_2 SKUs and assign them to forward pick area 2 and so on. The last $|N| - \sum_{m \in M} k_m$ SKUs by rank are allocated to the reserve. Let the corresponding assignment be $\mathbf{x}$.
>
> Determine optimal allocation, $\mathbf{v}$, for this assignment using 4.2.2
>
> For such an assignment, determine the objective value, $Z(\mathbf{x})$, from the expression (4.2.3a)
>
> Set $S \leftarrow \max\{S, Z(\mathbf{x})\}$

Return S and the corresponding assignment, $\mathbf{x}^*$, that generates the net benefit S. If $S = 0$, the assignment is $\mathbf{x}^* = 0$. Using this assignment, determine the allocations using (4.2.2).

Corollary 4.3.4. Algorithm 4 requires a computational effort of

$$O(|N| \log |N| + |M| \log |M| + |N|^{|M|}).$$

Proof. The time taken to sort the $|N|$ SKUs is $|N| \log |N|$ and the time to sort $|M|$ forward pick areas is $|M| \log |M|$. The number of possible $|M|$-tuples is $|N|^{|M|}$. □

Typically the number of forward pick areas in a warehouse is small and the computational effort of Algorithm 4 is hence equivalent to a low order polynomial in $|N|$. However, when the number of SKUs is large and the number of forward pick areas is small (say 5), this could lead to a large number of computations. To address this problem, Bartholdi and Hackman [6] analyze the continuous relaxation of the discrete problem and prove that the search can be executed much more efficiently due to a *sequential unimodality* property of the optimal value function. Therefore, the search need not be an exhaustive one over all $|N|^{|M|}$ possibilities, but instead, only needs to be a sequential bisection search that can be achieved in $O(\log^{|M|} |N|)$ time.

In the next section, we consider warehouses with limitations on labor and material flow and show how to extend the ranking methodology to assign and allocate space under such constraints.

4.4 Labor and material flow limits

Consider a warehouse with a conveyor from the forward pick area to the packaging location. The quantity picked, or flow, from the forward area per unit time is limited by the size and speed of the conveyor. Similarly, when a conveyor is used to restock the forward pick area, the flow of material is limited by the capacity of the conveyor. In addition, based on the number of workers or type of forward pick area, say Automatic Storage and Retrieval Systems (AS/RS) such as carousels, there are limitations on the number of picks and number of restocks that can be made per unit time.

We consider extensions to the fluid model where we include constraints that limit the number of restocks allowed by forward pick area or the total number of restocks across all forward pick areas and also constraints that limit the number of picks by forward pick area or across all forward pick areas. In addition, we include constraints on flow by forward pick area as well as across all forward pick areas.

Frazelle et al. [14] present an extension of Algorithm 3 to assign and allocate space in a warehouse with one forward pick area having constraints on *congestion*. They define the constraint on congestion as limiting the sum total of time spent in picking and restocking activities to a maximum allowable quantity. They use the fluid model to mathematically represent the picking time as proportional to the total picks from the forward pick area and the restocking time proportional to the total number of restocks required. Using the notation for the single mode, the constraint is

$$\sum_{i \in N} x_i \left(\alpha p_i + \beta \frac{f_i}{v_i} \right) \leq L, \tag{4.4.1}$$

where α and β are parameters of the warehouse based on layout. The right hand side, L, is the cap on total time for picking and restocking.

We consider a generalization of the problem, where in addition to the constraint on congestion, there are constraints on number of picks and number of restocks that is possible within the planning horizon. This is useful where the picking and restocking depend on type of equipment or amount of space available for movement of stock between fast pick area and reserve. For example, a conveyor based restocking system can only replenish the forward pick area at a certain rate. Therefore there is a natural limit for restocks that can be done within a time period. Similarly, constraints on picks could be due to pick lane size in the fast pick area, or, due to separate hours set aside for picking that limit the number of picks possible.

In the next section, we consider a warehouse with one forward pick area constrained by labor and material flow and show how to extend the ranking-based algorithm to determine a near-optimal assignment and allocation of SKUs.

4.5 Constraints in a warehouse with one forward pick area

We extend the fluid model for the warehouse with one forward pick area shown in (3.2.1) to include the constraints on congestion, picks and restocks as shown in formulation (4.5.1) below. The constraint in (4.5.1c) is the constraint on labor from Frazelle et al. [14] as defined in (4.4.1). The constraint (4.5.1d) limits the total number of restocks to R and the constraint (4.5.1e) caps the total number of picks to P during the planning horizon.

$$\max_{\mathbf{x},\mathbf{v}} \sum_{i \in N} x_i \left(s p_i - \frac{c_r f_i}{v_i} \right) \tag{4.5.1a}$$

subject to

$$\sum_{i\in N} x_i v_i \le V_F \tag{4.5.1b}$$

$$\sum_{i\in N} x_i \left(\alpha p_i + \beta \frac{f_i}{v_i}\right) \le L \tag{4.5.1c}$$

$$\sum_{i\in N} x_i \frac{f_i}{v_i} \le R \tag{4.5.1d}$$

$$\sum_{i\in N} x_i p_i \le P \tag{4.5.1e}$$

$$x_i \in \{0,1\} \qquad i \in N \tag{4.5.1f}$$

$$v_i \ge 0 \qquad i \in N \tag{4.5.1g}$$

Note that when $\mathbf{x}$ is fixed, the constraints on number of restocks (4.5.1d) are part of the objective function, and so they are also at the minimum possible value. If, for a given assignment of SKUs to forward pick areas, the restock constraints are infeasible then the assignment, $\mathbf{x}$, is also infeasible. Therefore, restating the formulation after eliminating $\mathbf{v}$ using (3.2.2), we get:

$$Z(x) \equiv \max_{\mathbf{x}} \left[\sum_{i\in N} x_i s p_i - \frac{c_r}{V_F}\left(\sum_{i\in N} x_i \sqrt{f_i}\right)^2\right] \tag{4.5.2a}$$

subject to

$$\sum_{i\in N} x_i \alpha p_i + \frac{\beta}{V_F}\left(\sum_{i\in N} x_i \sqrt{f_i}\right)^2 \le L \tag{4.5.2b}$$

$$\frac{1}{V_F}\left(\sum_{i\in N} x_i \sqrt{f_i}\right)^2 \le R \tag{4.5.2c}$$

$$\sum_{i\in N} x_i p_i \le P \tag{4.5.2d}$$

$$x_i \in \{0,1\} \qquad i \in N \tag{4.5.2e}$$

We assume the SKUs are ranked in descending order of labor efficiency, breaking ties arbitrarily. With this ranking, we show that a near-optimal solution can be obtained using an algorithm similar to Algorithm 3 presented in Chapter 3.

Theorem 4.5.1. There exists an optimal assignment $\tilde{\mathbf{x}}$ to the linear relaxation of problem (4.5.2) such that SKUs of a higher rank have a higher claim to the forward pick area. That is if SKU i outranks j, then, $\bar{x}_i \geq \tilde{x}_j$ and if $\tilde{x}_j > 0$, then $\tilde{x}_i = 1$.

Proof. We take a similar approach to the proof as in Theorem 3.3.1. Let $\tilde{\mathbf{x}}$ be an optimal solution to linear relaxation of problem (4.5.2). If the optimal solution $\bar{\mathbf{x}}$, has the properties we seek, then we are done. If not, then there exists two SKUs i and j where, i outranks j and $\tilde{x}_j > \tilde{x}_i$. We modify this solution by switching SKUs to construct an alternate optimal solution with the desired properties.

Construct another solution, $\bar{\mathbf{x}}$,

$$\bar{x}_k = \begin{cases} \tilde{x}_k & k \in N \setminus \{i, j\} \\ \tilde{x}_i + \epsilon p_j / p_i & k = i \\ \tilde{x}_j - \epsilon & k = j \end{cases}$$

We choose the maximum possible $\epsilon > 0$ such that $\bar{\mathbf{x}}$ is a feasible solution. That is, either $\bar{x}_i = 1$ or $\bar{x}_j = 0$ or both. Such an ϵ exists because $\tilde{x}_i < 1$. Note that by choice of ϵ, $\bar{x}_i = 1$ if $\bar{x}_j > 0$. Rewriting the objective function,

$$Z(\bar{\mathbf{x}}) = \sum_{k \in N} \tilde{x}_k s p_k - \frac{c_r}{V_F} \left[\sum_{k \in N} \tilde{x}_k \sqrt{f_k} + \epsilon \left(\sqrt{f_i} p_j / p_i - \sqrt{f_j} \right) \right]^2 .$$

Note that the pick savings remain the same and only the restocking costs change. Since i outranks j, that is, labor efficiency of i is at least or greater than that of j, and so,

$$\frac{p_i}{\sqrt{f_i}} \geq \frac{p_j}{\sqrt{f_j}}.$$

Or,

$$\sqrt{f_i} p_j / p_i - p_j \leq 0.$$

Therefore, the restocking costs remain the same or decrease. This implies that $Z(\bar{\mathbf{x}}) \geq Z(\tilde{\mathbf{x}})$ and also that the constraint (4.5.2c) is still feasible. The constraint on picks, (4.5.2d) is also feasible since the pick savings does not change. This also implies that the constraint on congestion, (4.5.2b), is also feasible since the total number of picks is the same and the number of restocks remain the same or decrease than for the optimal solution. Similarly, other pairs of SKUs not in rank order can be switched to construct a solution with the desired properties while maintaining feasibility and optimality. □

We see that from Theorem 4.5.1 that the higher ranked SKUs are allocated first. This immediately implies that there is an optimal solution to the linear relaxation such that at most one SKU i is fractional, i.e. $\bar{x}_i \in (0, 1)$.

Corollary 4.5.2. The optimal solution to the linear relaxation of problem (4.5.2) has at most one fractional SKU.

Proof. From Theorem 4.5.1. □

We present an adaptation of Algorithm 3 to generate a solution that is a close approximation of the optimal solution to problem (4.5.2).

Corollary 4.5.3. Algorithm 5 generates a solution that is within the net benefit of one SKU from optimal and requires a computational effort of at most $|N| \log |N| + |N|$

Proof. From Theorem 4.5.1, we see that there is an optimal solution, where the SKUs of highest labor efficiency with the corresponding $x > 0$. If there is a fractional SKU, there is at most one and is the least ranked SKU. The Algorithm 5 produces a solution that has all but the fractional SKU in common with the optimal. Therefore it is within at most the net benefit of one SKU from optimal.

The time taken to sort the $|N|$ SKUs is $|N| \log |N|$ and the loop in step 3 of Algorithm 5 runs $|N|$ times. □

In this case too, the algorithm to find a near-optimal solution uses an a priori ranking for SKUs similar to the case of the warehouse without constraints. The algorithm ranks the SKUs in the order of labor efficiency and assigns SKUs to the

Algorithm 5: Algorithm to determine SKU assignment and allocations for a warehouse with one forward pick area that is constrained by limits on congestion, number of picks and number of restocks required.

Sort the SKUs in the set N in the descending order of their labor efficiencies, breaking ties arbitrarily
Set $S \leftarrow 0$
for $k = 1$ **to** $|N|$ **do**
  Select the first k SKUs and assign them to the forward pick area F, i.e., set their respective x to 1 and the rest to 0.
  If any of the constraints (4.5.2b), (4.5.2d) or (4.5.2c) are violated, exit the loop.
  Determine optimal allocation, $\mathbf{v}$, for this assignment using 3.2.2
  For assignment $\mathbf{x}$, determine the objective value, $Z(\mathbf{x})$, from the expression (4.5.2a)
  if $\boldsymbol{x}$ *and* $\boldsymbol{v}$ *satisfy the constraints (4.5.1c)-(4.5.1e)* **then**
    if $S < Z(\boldsymbol{x})$ **then**
      Set $\bar{\mathbf{x}} = \mathbf{x}$
      Set $S \leftarrow Z(\mathbf{x})$
return assignment $\bar{\mathbf{x}}$.

forward pick areas from the most desired to the least until one of the constraints is maximally satisfied. The solution generated by the algorithm is near-optimal, within the net benefit of a single SKU from optimal. Given that warehouses have hundreds, if not, thousands of SKUs, the bound is very tight in practice. This methodology can be extended to warehouses with constraints on multiple forward pick areas as we show in the next section.

4.6 Constraints in a warehouse with multiple forward pick areas

In the case of a warehouse with multiple forward pick areas, we consider constraints that impose limits on labor and material flow similar to the single forward pick area. We impose the congestion constraint similar to (4.4.1) that limits congestion across all forward pick areas, where the parameter α remains the same for all forward pick areas and the parameter β_m can vary by the forward pick area m. In addition, we also have a constraint on the number of restocks, individually for each forward pick area as well as in aggregate. Lastly, we also consider a constraint on total number of picks from the forward pick areas in aggregate. All these constraints impose different types of limits on use of labor or material flow that warehouses face. For example, the constraints across all forward pick areas for restocks and picks can enforce limitations based on common resources like order pickers, conveyor belts that maybe used to transport the SKUs, restocking time available or aisle widths. We extend the formulation (4.2.1) to include these additional constraints as shown in (4.6.1).

$$\max_{\mathbf{x},\mathbf{v}} \sum_{m\in M}\sum_{i\in N} x_{mi}\left(s_m p_i - \frac{c_{rm} f_i}{v_{mi}}\right) \tag{4.6.1a}$$

subject to

$$\sum_{i\in N} x_{mi} v_{mi} \le V_m \qquad \forall m \in M \tag{4.6.1b}$$

$$\sum_{m\in M} x_{mi} = 1 \qquad \forall i \in N, \tag{4.6.1c}$$

$$\sum_{i\in N} x_{mi}\frac{f_i}{v_{mi}} \le R_m \qquad \forall m \in M \tag{4.6.1d}$$

$$\sum_{\substack{i\in N\\ m\in M}} x_{mi}\frac{f_i}{v_{mi}} \le R \tag{4.6.1e}$$

$$\sum_{\substack{i\in N\\ m\in M}} x_{mi} p_i \le P \tag{4.6.1f}$$

$$\sum_{\substack{i\in N\\ m\in M}} x_{mi}\left(\alpha p_i + \beta_m \frac{f_i}{v_{mi}}\right) \le L \tag{4.6.1g}$$

$$x_{mi} \in \{0,1\} \qquad m \in M, i \in N, \tag{4.6.1h}$$

$$v_{mi} \ge 0 \qquad m \in M, i \in N \tag{4.6.1i}$$

The constraint (4.6.1d) imposes a cap on the number of restocks for each forward pick area and (4.6.1e) does the same on the aggregate number of restocks. The constraint on total number of picks from all forward pick areas and congestion is represented in (4.6.1f) and (4.6.1g). Note that if there is only one forward pick area, then the aggregate constraints across all forward pick areas, (4.6.1e) and (4.6.1f), are in effect, equivalent to the single forward pick area constraints.

Note that when $\mathbf{x}$ is fixed, the constraints representing number of restocks in (4.6.1e), (4.6.1d) and (4.6.1g) are part of the objective function, and so they are also at the minimum possible value. Therefore, we can represent them using the optimal solution for the allocations, $\tilde{\mathbf{v}}$ shown in (4.2.2). This is so because of the value of $\mathbf{x}$, the constraints are infeasible, then the assignment $\mathbf{x}$ is an infeasible solution.

Eliminating $\mathbf{v}$ in (4.6.1) using expression (4.2.2), we get the equivalent formulation,

$$Z(x) \equiv \max_{\mathbf{x},\mathbf{v}} \sum_{m \in M} \left[\sum_{i \in N} x_{mi} s_m p_i - \frac{c_{rm}}{V_m} \left(\sum_{i \in N} x_{mi} \sqrt{f_i} \right)^2 \right] \tag{4.6.2a}$$

subject to

$$\sum_{m \in M} x_{mi} = 1 \qquad \forall i \in N, \tag{4.6.2b}$$

$$\frac{1}{V_m} \left(\sum_{i \in N} x_{mi} \sqrt{f_i} \right)^2 \leq R_m \qquad \forall m \in M \tag{4.6.2c}$$

$$\sum_{m \in M} \frac{1}{V_m} \left(\sum_{i \in N} x_{mi} \sqrt{f_i} \right)^2 \leq R \tag{4.6.2d}$$

$$\sum_{\substack{i \in N \\ m \in M}} x_{mi} p_i \leq P \tag{4.6.2e}$$

$$\sum_{\substack{i \in N \\ m \in M}} x_{mi} \alpha p_i + \sum_{m \in M} \beta_m \frac{1}{V_m} \left(\sum_{i \in N} x_{mi} \sqrt{f_i} \right)^2 \leq L \tag{4.6.2f}$$

$$x_{mi} \in \{0, 1\} \qquad m \in M, i \in N \tag{4.6.2g}$$

A result similar to Theorem 4.3.1 holds for the linear relaxation of the formulation (4.6.2), where SKUs of a higher rank have a higher claim on the higher ranked forward pick areas.

Theorem 4.6.1. Let $m, w \in M$, be two forward pick areas such that m outranks w. Then there is an optimal solution, $\tilde{\mathbf{x}}$, where for two SKUs $i, j \in N$ such that $\tilde{x}_{mi} > 0$

and $\tilde{x}_{wj} > 0$, SKU i outranks SKU j. Further, in that solution, at least one of $\tilde{x}_{mj}$ and $\tilde{x}_{wi}$ are zero.

Proof. We proceed in a manner very similar to the proof of Theorem 4.3.1 but with a minor difference. We assume that there does not exist such a solution. Let $\tilde{\mathbf{x}}$ be an optimal solution, where, for two forward pick areas $m, w \in M$, two SKUs $i, j \in N$, m outranks w, $\tilde{x}_{mj} > 0$ and $\tilde{x}_{wi} \geq 0$ and i outranks j. By modifying the optimal solution, we construct alternative solutions that satisfy the required properties and contradict the assumption.

Consider the case where $\tilde{x}_{wi} = 0$. This implies there is no forward pick area w where SKU i is stored, or equivalently, SKU i which is higher ranked than SKU j is assigned to the reserve. We construct an alternate feasible solution $\bar{\mathbf{x}}$ by modifying $\tilde{\mathbf{x}}$ similar to Theorem 4.3.1 but in a slightly different manner:

$$\bar{x}_{mk} = \begin{cases} \tilde{x}_{mk} & k \in N \setminus \{i, j\} \\ \tilde{x}_{mi} + \epsilon p_j / p_i & k = i \\ \tilde{x}_{mj} - \epsilon & k = j \end{cases}$$

As shown in Theorem 3.3.1, this solution is feasible and also has the same pick savings as the optimal assignment $\tilde{\mathbf{x}}$. In addition, the restocking cost is at most the corresponding value for the optimal solution as $p_i/\sqrt{f_i} \geq p_j/\sqrt{f_j}$. Therefore the total net benefit of the new assignment $\bar{\mathbf{x}}$ is:

$$\begin{aligned} Z(\bar{\mathbf{x}}) = & \sum_{l \in M, k \in N} \tilde{x}_{lk} s p_k - \sum_{l \in M \setminus \{m\}} \frac{c_r}{V_l} \left(\sum_{k \in N} \tilde{x}_{lk} \sqrt{f_k} \right)^2 \\ & - \frac{c_r}{V_m} \left[\sum_{k \in N} \tilde{x}_{mk} \sqrt{f_k} + \epsilon \left(\sqrt{f_i} p_j / p_i - \sqrt{f_j} \right) \right]^2 . \end{aligned}$$

Following Theorem 4.3.1, we choose the maximum possible value for ϵ and therefore, either $(\bar{x}_{mi}, \bar{x}_{mj})$ is $(1, \bar{x}_{mj})$ or $(\bar{x}_{mi}, 0)$. The constraints on picks are still feasible as the number of picks has not changed. Also feasible are the constraints on number of restocks and congestion since the number of restocks has only reduced or remained the same for all forward pick areas. This contradicts our assumption that no such solution exists. Therefore, it must be that $\tilde{x}_{wi} > 0$ for some forward pick area w that is outranked by forward pick area m.

The rest of the proof follows in a manner similar to Theorem 4.3.1, where we construct an alternative solution, $\bar{\mathbf{x}}$, with the same restocking cost but with a pick savings that is at least as much as the optimal value. The constraints on number of restocks are still feasible for the assignment $\bar{\mathbf{x}}$ since the restocking cost is the same as the optimal for all forward pick areas. Also the constraints involving the number of picks is also satisfied, since the total number of picks from the forward pick areas has not changed but reassigned from one forward pick area to another. Hence the contradiction to the assumption that no optimal solution exists with the desired properties. $\square$

Therefore, Corollary 4.3.2 immediately follows and a simple adaptation of Algorithm 4 shown below in Algorithm 6 constructs a solution, $\bar{\mathbf{x}}$, that is near-optimal with the same complexity. Note that for any given assignment $\mathbf{x}$, the space allocation is estimated using the expression (4.2.2).

The solution, $\bar{\mathbf{x}}$, is within the net benefit of $|M|$-SKUs from optimal and since the number of SKUs is large and number of forward pick areas is small in a typical warehouse, this is a very good practical bound. The computational complexity of the algorithm is identical to that of Algorithm 4, which produces a near-optimal assignment and allocation of SKUs for the case of the warehouse with multiple forward pick areas and no constraints. In addition, the sequential unimodality property shown by Bartholdi and Hackman [23] for the continuous relaxation of the optimal function value is also valid here and therefore, the complexity of the Algorithm 6 can be similarly improved.

The immediate extension is the more general case of determining the optimal assignment and allocation of SKUs for warehouses with multiple forward pick areas

Algorithm 6: Algorithm to determine assignment and allocation of SKUs for a warehouse with multiple forward pick areas that are constrained by limits on congestion, total number of picks and number of restocks.

Sort the SKUs in the set N in the descending order of their labor efficiencies, breaking ties arbitrarily
Sort the forward pick areas in the set M in the descending order of their pick savings, breaking ties arbitrarily. Let the forward pick areas be $M = \{1, 2, \dots, |M|\}$ and since the pick savings of the reserve is zero as well as the least, it would be the last forward pick area in the set.
Set $S \leftarrow 0$
foreach $\mathbf{k} = (k_1, k_2, \dots, k_{|M|})$ *ordered* $|M|$*-tuple such that* $\sum_{m \in M} k_m = |N|$ *and* $k_m \in \{0, 1, \dots, |N|\}$ **do**
    Determine an assignment vector $\mathbf{x}$ such that the first k_1 SKUs are assigned to forward pick area 1, k_2 SKUs to forward pick area 2 and so on. That is, set the corresponding x to 1 and the rest to 0.
    Determine optimal allocation, $\mathbf{v}$, for this assignment using (4.2.2)
    if $\boldsymbol{x}$ *and* $\boldsymbol{v}$ *satisfy the constraints (4.6.1d)-(4.6.1g)* **then**
        For assignment $\mathbf{x}$, determine the objective value $Z(\mathbf{x})$ using expression (4.6.2a)
        if $S < Z(\boldsymbol{x})$ **then**
            Set $\bar{\mathbf{x}} = \mathbf{x}$
            Set $S \leftarrow Z(\mathbf{x})$
return assignment $\bar{\mathbf{x}}$.

that have individual constraints on number of picks by forward pick area in addition to the constraints on number of restocks and congestion. This allows restrictions due to layout, picking equipment and labor limits to be placed in a similar manner on both quantities, the number of picks and the number of restocks. The formulation of this problem is shown in (4.6.3).

$$\max_{\mathbf{x},\mathbf{v}} \sum_{m \in M} \sum_{i \in N} x_{mi} \left(s_m p_i - \frac{c_{rm} f_i}{v_{mi}} \right) \tag{4.6.3a}$$

subject to

$$\sum_{i \in N} x_{mi} v_{mi} \leq V_m \qquad \forall m \in M \tag{4.6.3b}$$

$$\sum_{m \in M} x_{mi} = 1 \qquad \forall i \in N, \tag{4.6.3c}$$

$$\sum_{i \in N} x_{mi} \frac{f_i}{v_{mi}} \leq R_m \qquad \forall m \in M \tag{4.6.3d}$$

$$\sum_{\substack{i \in N \\ m \in M}} x_{mi} \frac{f_i}{v_{mi}} \leq R \tag{4.6.3e}$$

$$\sum_{\substack{i \in N \\ m \in M}} x_{mi} p_i \leq P \tag{4.6.3f}$$

$$\sum_{i \in N} x_{mi} p_i \leq P_m \qquad \forall m \in M \tag{4.6.3g}$$

$$\sum_{\substack{i \in N \\ m \in M}} x_{mi} \left(\alpha_m p_i + \beta_m \frac{f_i}{v_{mi}} \right) \leq L \tag{4.6.3h}$$

$$x_{mi} \in \{0, 1\}, v_{mi} \geq 0 \qquad m \in M, i \in N \tag{4.6.3i}$$

The constraints (4.6.3g) enforces a limit on number of picks by forward pick area. The constraint on congestion, (4.6.3h), also uses the parameter α_m to vary the effect on congestion by forward pick area m. This allows individual modeling of constraints related to picking for each of the forward pick areas.

Though it may appear that the ranking-based heuristic from Algorithm 6 can be extended to solve this formulation as well, the individual constraints on number of picks by forward pick area seem not to allow it. The switching based arguments that we used extensively in earlier cases do not apply here because whenever the constraint on number of picks is tight, we cannot switch a lower rank SKU with a higher rank one as that would violate the constraint. To illustrate that the performance of ranking-based algorithm in the case of a warehouse with multiple forward pick areas with constraints on picks by forward pick area, we construct an example data set where it is much further from the optimal than the net benefit of one SKU.

Consider a warehouse with two forward pick areas F_1 and F_2. Table 1 lists for each forward pick area the space available, cost parameters and caps on number of picks and restocks that can be done within the planning horizon. The warehouse

Table 1: Parameters of a warehouse with two forward pick areas where Volume is the space available, Pick Savings is the savings in cost per pick compared to the reserve, Restock Cost is the cost per restock, Max Picks and Max Restocks are the maximum number of picks and restocks possible during the planning horizon.

Forward Pick Area	Volume	Pick Savings	Restock Cost	Picks	Restocks
F_1	10,000	4.0	5	15,000	3,000
F_2	10,000	3.9	5	30,000	1,500

has 1000 SKUs that are broken into two classes of 500 each. In Table 2 we show for each class has the expected number of picks and flow during the planning horizon. Observe that on the basis of pick savings, forward pick area F_1 outranks forward

Table 2: The number of SKUs, expected number of picks and flow per SKU during the planning horizon for the two SKU classes in the warehouse.

Class	SKUs	Picks	Flow
A	500	60	60
B	500	30	120

pick area F_2. And on the basis of labor efficiency, SKUs in class A outrank those in class B. Using a simple adaptation of Algorithm 6 to assign and allocate SKUs by enumerating all solutions such that the higher ranked forward pick areas area are assigned higher ranked SKUs within the constraints of formulation (4.6.3), we get the solution shown in Table 3. This solution has a total net benefit of 129,749.916. However, it is exceeded by over 19% by the net benefit of the solution that stocks the SKUs in the reverse order of labor efficiency as shown in Table 4. The optimal net benefit is at least as much the solution in Table 4 and therefore, the ranking-based solution is at least 16% less than the optimal. The net benefit of any one SKU does

Table 3: An assignment to the forward pick areas using an adaptation of Algorithm 6 to assign and allocating SKUs by enumerating all solutions such that the higher ranked forward pick areas are assigned higher ranked SKUs within the constraints on number of picks and restocks.

Forward Pick Area	SKUs of Class A	SKUs of Class B
F_1	250	250
F_2	0	176

Table 4: An alternative assignment to the forward pick areas that assigns the SKUs in reverse order of their rank. The total net benefit of the ranking-based solution in Table 3 is less than the corresponding value for this solution by 16%.

Forward Pick Area	SKUs of Class A	SKUs of Class B
F_1	0	500
F_2	500	0

not exceed 0.2% of the total net benefit, and 16% is hence a significant difference even if the gap is represented in terms of net benefit of a certain number of SKUs. Though ranking-based algorithms generate solutions that are near-optimal for the extensions considered until now, they do not in this case.

4.7 Conclusions

In Chapter 3, we present an alternative argument to prove the effectiveness of the greedy heuristic developed by Hackman and Rosenblatt [23] for a warehouse with one forward pick area. In this chapter, we show how to extend the same argument to show ranking heuristic proposed by Bartholdi and Hackman [6] for a warehouse with multiple forward pick areas is near-optimal. The SKUs are ranked as before by their labor efficiencies and the forward pick areas by their pick savings. The near-optimal solution generated by the algorithm has the property that the higher ranked forward pick areas have a higher claim to the high rank SKUs. This allows the warehouse manager to sort the SKUs and forward pick areas and also easily re-slot SKUs as the assortment carried by the warehouse evolves.

We extend the ranking methodology to warehouses with constraints on congestion, total number of picks and number of restocks, in total and by forward pick area. These constraints occur due to limited availability of the workforce and any scheduling restrictions that may apply. Also the warehouse layout and equipment employed like carousels or conveyor belts can also limit movement of SKUs. We use the same proof technique to show a similar ranking algorithm with a minor modification to accommodate the constraints has the same near-optimality bounds as well as computational efficiencies. This allows the warehouse manager to take advantage of the ranking framework for forward pick areas with constraints as well. However, in the case of the warehouse with multiple forward pick areas with individual limits on the number of picks using we see that the ranking-based algorithm does not obtain a near-optimal solution indicating that this methodology has its limitations.

Until now, once the SKUs are assigned, we assume that the space in the forward pick area is allocated based on the optimal allocation shown in (4.2.2). However, most warehouses use more conventional strategies of allocating space, such as Equal Space (equal space for all SKUs assigned) or Equal Time (SKUs assigned have an equal time supply). In the next chapter, we discuss how to assign SKUs to forward pick areas with either stocking strategy such that total cost is minimized.

CHAPTER V

ASSIGNING SKUS IN A WAREHOUSE USING EQUAL SPACE AND EQUAL TIME STRATEGIES

5.1 Allocating space in other ways

Most warehouses, instead of using optimal space allocations, prefer to allocate space in conventional ways: the assigned SKUs are either equal in space, or equal in time to restock [8]. We refer to these two types of stocking strategies as Equal Space and Equal Time respectively. In this chapter, we show how to assign SKUs to forward pick areas for such warehouses that use Equal Space or Equal Time to allocate space, so as to minimize the sum of picking and restocking costs. We construct polynomial algorithms to do so, making it practically efficient to determine the assignments of SKUs for warehouses with one or more forward pick areas. We test the implementation of these algorithms with data from three warehouses of different sizes and compare the best solution and run time of the algorithms.

We also look at warehouses employing Equal Space and Equal Time allocations that have constraints due to limitations on material flow and labor availability. We represent these constraints in a manner similar to the discussion in Chapter 4 as a cap on number of picks and number of restocks for each forward pick area.

5.2 A warehouse with one forward pick area using Equal Space

In Chapter 3, we presented an algorithm that determines a near-optimal assignment and space allocation of SKUs to forward pick areas such that labor cost is minimized. In the optimal strategy, given any assignment, we determine the space allocation so as to minimize restocking costs. An alternate allocation strategy is to allocate equal space to each SKU. We extend the methodology to handle this case where the SKUs assigned to forward pick areas share the space equally. In this section, we address the question of which SKUs should be assigned to the forward pick area so as to minimize total cost given the equal space storage policy for the chosen SKUs.

We extend the Hackman and Rosenblatt [23] fluid model to accommodate the equal space allocation. If SKUs were allocated equal space, we see that the space allocated to SKU i assigned to the forward pick areas is $v_i = V_F / \sum_{j \in N} x_j$, where, $\sum_{j \in N} x_j$ is the number of SKUs allocated to the forward pick area. Using this observation in the model in (3.2.1), and substituting for $\mathbf{v}$, we get the model in (5.2.1).

$$\max_{\mathbf{x}} \sum_{i \in N} x_i \left(sp_i - \frac{c_r f_i \sum_{j \in N} x_j}{V_F} \right) \tag{5.2.1a}$$

subject to

$$x_i \in \{0, 1\} \qquad i \in N \tag{5.2.1b}$$

Let the number of SKUs assigned to the forward pick area F be $n_F = \sum_{j \in N} x_j$. Each SKU i in those n_F SKUs adds $sp_i - \frac{c_r f_i n_F}{V_F}$ in net benefit. To maximize total net benefit, we can determine these net benefit for all SKUs and choose the top n_F SKUs by ordering them in descending order of net benefit. That would give the optimal allocation for n_F SKUs and helps construct Algorithm 7.

Algorithm 7: Algorithm to determine optimal assignment for a warehouse with one forward pick area using Equal Space allocation.

---

for $n_F = 1$ **to** $|N|$ **do**
- For each SKU i, compute net benefit as $z_i = sp_i - \frac{c_r f_i n_F}{V_F}$.
- Sort z_i in descending order and choose the first n_F SKUs. This is the best allocation for n_F SKUs
- **if** *any SKU has a negative net benefit* **then** break
- Determine the total net benefit, $Z(n_F)$, as the sum of z_i for the top n_F SKUs

Choose the allocation, or n_F, that produces the maximum $Z(n_F)$

We show this algorithm to be computationally efficient. Just as in the labor efficient allocation, by enumerating only $|N|$ possibilities, we can determine the optimal allocation.

Theorem 5.2.1. Algorithm 7 produces an optimal equal space allocation as defined in formulation (5.2.1) in $O(|N|^2 \log |N|)$ time.

Proof. We see that in problem formulation (5.2.1), setting $n = \sum_{j \in N} x_j$ causes the objective function to separate. Hence, the function can be maximized by choosing the top n values of z_i as in step 3 of Algorithm 7. The possible values for n_F range from 1 to $|N|$ and determining the objective for all possible values of n_F, we get the optimal allocation.

Also note that when n_F is large enough, it maybe that the least z_i values of the top n_F SKUs could be negative. In such cases, any further increase in n_F will only make them further negative and the additional SKUs introduced will have even lower negative values. Therefore the first occurrence of a negative net benefit is a termination condition when enumerating the values of n_F.

The number of possible values for n_F is $|N|$ and for each n_F, the effort to calculate z_i and sort is $|N| + |N| \log |N|$. Hence the complexity of the algorithm is $O(|N|^2 \log |N|)$. □

The algorithm to determine the optimal solution is computationally efficient, but, it does not offer the additional insights of the corresponding ranking-based algorithm when determining the best assignment with optimal space allocations. As discussed in Chapter 3, the latter provides an a priori ranking of SKUs that indicates the claim of a SKU to the forward pick area that warehouse managers can take advantage of. This would mean that when using the Equal Space allocation and minor changes in system parameters occur like new SKUs added, a few existing ones removed or, space in the forward pick area changed, there maybe a completely different optimal assignment of SKUs instead of a modification on the existing assignment.

In addition, there are no unimodality properties that the algorithm can utilize when determining the best Equal Space allocation. This is unlike the labor efficient case, where Bartholdi and Hackman [6] demonstrate unimodality in the search for the best allocation. Therefore all $|N|$ possibilities have to be evaluated in the search for the optimal solution. The following example illustrates a case where Algorithm 7 demonstrates bimodal behavior where the net benefit is shown to increase and decrease twice.

Consider a warehouse with the parameters shown in Table 5. Note that the Algorithm evaluates $sp_i - \frac{c_r f_i n_F}{V_F}$ at each step for different SKUs and different values of n_F. Therefore, for ease of exposition, only the values necessary for applying Algorithm 7 are shown.

Table 5: An example warehouse with one forward pick area and 6 SKUs that demonstrates bimodal behavior for the sequential evaluation of objective function of total net benefit in Algorithm 7.

SKU	sp	$\frac{c_r f}{V_F}$
A	1.7	0.7
B	1.7	0.7
C	0.30	0.01
D	0.30	0.01
E	0.30	0.01
F	0.30	0.01

Running the algorithm for the data shown in Table 5, we evaluate the net benefit in each iteration for SKUs assigned to the forward pick area. The values of the net benefit as evaluated at each value of n_F is bimodal as can be seen from the Graph 11.

This example shows that that the Algorithm 7 is not unimodal and therefore the search may have to exhaustive over all values of n_F. In addition, note that the upward trend starting at $n_F = 3$ can be extended by adding more copies of SKU C.

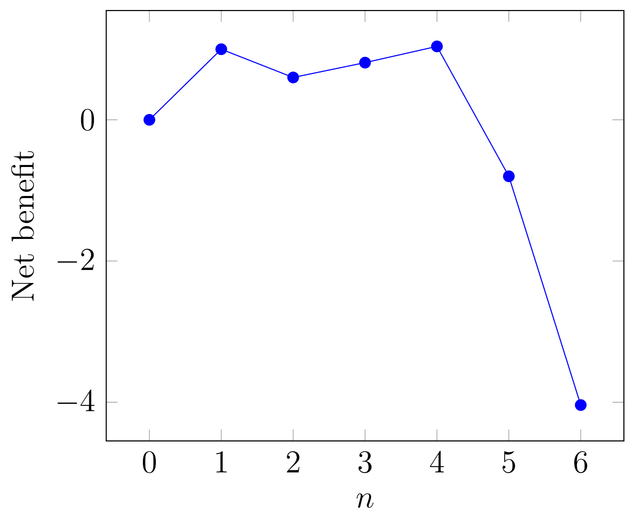

Figure 11: The net benefit of the SKUs assigned in each iteration for the warehouse whose parameters are shown in Table 5. The bimodal behavior of the plot shows that the Algorithm 7 has to be evaluated for every feasible n_F to determine the best solution.

For $3 \leq n_F \leq |N| - 2$, the net benefit would turn out to be $n_F(sp_C - n_F c_r f_C / V_F)$, which in this case is $n_F(0.3 - 0.01n_F)$. As long as $0.3 > 0.01n_F$, the trend would be increasing with n_F. Therefore, without assumptions on the data set, the complexity of the algorithm cannot be reduced.

5.3 A warehouse with multiple forward pick areas using Equal Space

Warehouses that have multiple forward pick areas with different picking and restocking cost face an additional challenge of choosing which forward pick area to use for which SKU. With $|M|$ forward pick areas present in the warehouse, the fluid model formulation is extended as follows in (5.3.1). We extend the notation to address multiple forward pick areas with V_l, s_l and c_{rl} representing the volume, picking cost savings and restock cost of forward pick area l respectively. The decision variable x_{li} is set to 1 if SKU i is allocated to forward pick area l. As before N is the set of SKUs and $M = \{1, 2, \ldots, |M|\}$ is the set of forward pick areas.

$$\max_{\mathbf{x}} \sum_{i \in N, l \in M} x_{li} \left(s_k p_i - \frac{c_{rl} f_i \sum_{j \in N} x_{lj}}{V_l} \right) \tag{5.3.1a}$$

subject to

$$\sum_{l \in M} x_{li} \leq 1 \qquad i \in N \tag{5.3.1b}$$

$$x_{li} \in \{0, 1\} \qquad i \in N, l \in M \cup \{0\} \tag{5.3.1c}$$

We rewrite the formulation (5.3.1) as an equivalent minimization problem using picking and restocking costs for each forward pick area and the reserve. We group the forward pick areas M and the reserve, represented as 0, as *modes* and represent the set as $M \cup \{0\}$. We use the notation d_l to denote picking cost from mode l where $l \in M \cup \{0\}$. The restock cost, c_{r0}, is zero for the reserve.

$$\min_{\mathbf{x}} \sum_{i \in N, l \in M \cup \{0\}} x_{li} \left(d_l p_i + \frac{c_{rl} f_i \sum_{j \in N} x_{lj}}{V_l} \right) \tag{5.3.2a}$$

subject to

$$\sum_{l \in M \cup \{0\}} x_{li} = 1 \qquad i \in N \tag{5.3.2b}$$

$$x_{li} \in \{0, 1\} \qquad i \in N, l \in M \cup \{0\} \tag{5.3.2c}$$

In the case of the single forward pick area, Algorithm 7 enumerated over the possible number of SKUs in the forward pick area and determined the optimal solution given the number of SKUs, n, to be stored in the forward pick area. We then chose the number of SKUs that gave us the highest objective, or, benefit. We take the same approach in the case of multiple forward pick areas, enumerating over all possibilities of number of SKUs $\mathbf{n} = (n_0, n_1, n_2, \ldots, n_{|M|})$ assigned to the modes, determining the best possible solution for each combination and choosing the one with the lowest cost. Note that $\sum_{l \in M \cup \{0\}} n_l = |N|$.

Given the number of SKUs to be assigned to the modes, $\mathbf{n} = (n_0, n_1, \ldots, n_{|M|})$, the problem reduces to (5.3.3)

$$C(\mathbf{n}) = \min_{\mathbf{x}} \sum_{i \in N, l \in M \cup \{0\}} x_{li}(d_l p_i + \frac{c_{rl} f_i n_l}{V_l}) \tag{5.3.3a}$$

subject to

$$\sum_{l \in M \cup \{0\}} x_{li} = 1 \qquad i \in N \tag{5.3.3b}$$

$$\sum_{i \in N} x_{li} = n_l \qquad l \in M \cup \{0\} \tag{5.3.3c}$$

$$x_{li} \in \{0, 1\} \qquad i \in N, l \in M \cup \{0\} \tag{5.3.3d}$$

Theorem 5.3.1. The formulation (5.3.3) for each $\mathbf{n}$ can be solved in strongly polynomial time.

Proof. We can reduce the problem (5.3.3) to a capacitated min-cost flow problem, if we create a directed network graph $G(V, E)$ as follows. Set the vertex set $V = N \cup M \cup \{0\}$, and the directed edge set as $E = N \times (M \cup \{0\})$, with edge capacities as 1. Assign the cost of edge $e = (i, l)$, where $i \in N$ and $l \in M \cup \{0\}$, to be the cost of picking SKU i from mode l, which is $x_{li}\left(d_l p_i + \frac{c_{rl} f_i n_l}{V_l}\right)$. The supply at vertex $i \in N$ is 1 and at $l \in M \cup \{0\}$ is n_l. This min-cost flow problem is the linear relaxation of formulation (5.3.3) above.

$$\min_{\mathbf{x}} \sum_{i \in N, l \in M \cup \{0\}} x_{li}\left(d_l p_i + \frac{c_{rl} f_i n_l}{V_l}\right) \tag{5.3.4a}$$

subject to

$$\sum_{l \in M \cup \{0\}} x_{li} = 1 \qquad i \in N \tag{5.3.4b}$$

$$\sum_{i \in N} x_{li} = n_l \qquad l \in M \cup \{0\} \tag{5.3.4c}$$

$$x_{li} \geq 0 \qquad i \in N, l \in M \cup \{0\} \tag{5.3.4d}$$

The constraint matrix in formulation 5.3.4 is totally unimodular and hence the solution is integral. There are strongly polynomial algorithms to solve min-cost flows and one in particular that solves the problem in $O(|E| \log |V| (|E| + |V| \log |V|))$ time [29]. $\square$

Using the definition of E and V in Theorem 5.3.1, we can compute the complexity of solving formulation (5.3.3) using Orlin's algorithm [29]. Since $|E| = |N| (|M| + 1)$ and $|V| = |N| + |M| + 1$,

$$O(|E| \log |V| (|E| + |V| \log |V|)) \tag{5.3.5}$$

$$= O(|N| |M| \log(|N| + |M|)[|N| |M| + (|N| + |M|) \log(|N| + |M|)]) \tag{5.3.6}$$

$$< O(|N|^2 |M|^2 \log^2 |N|). \tag{5.3.7}$$

The last line follows because the number of SKUs is typically much greater than the number of forward pick areas, or, $|N| >> |M|$.

Therefore the algorithm for Equal Space Allocation in the case of multiple forward pick areas is as follows.

Algorithm 8: Algorithm to determine optimal assignment for a warehouse with multiple forward pick areas using Equal Space allocation.

foreach $\mathbf{n} = (n_0, n_1, n_2, \ldots, n_{|M|})$ $s.t$ $\sum_{l \in M \cup \{0\}} n_l = |N|$ **do**
    Evaluate optimal objective, $C(\mathbf{n})$, and allocation, $\mathbf{x}$, by solving minimum cost network flow problem (5.3.4)
Choose the allocation, $\mathbf{x}$, that produces the minimum cost $C(\mathbf{n})$

Since there are $|N|^{|M|}$ possible values of $\mathbf{n}$ and the min-cost flow has to be solved for each case, we get a net complexity of $O(|N|^{|M|+2} |M|^2 \log^2 |N|)$. In real warehouses M could be as high as 5, resulting in a rather high $O(|N|^7)$ running time. We address this issue by suggesting ways to cut down on the number of enumerations required to be evaluated in the next section.

Unlike the algorithms presented for labor efficient allocation and the equal space allocation for one forward pick area, we do not have a algorithm based on sequential objective evaluation using ranking properties. In this case, a network flow algorithm is required that is less intuitive than the other case, but solves the problem efficiently, but involves a more expensive operation of solving network flow problems at each iteration.

5.4 Improving the computational time

The algorithm to determining the optimal assignment in a warehouse with multiple forward pick areas using Equal Space allocation involves two time consuming steps:

1. The number of assignment combinations of $\mathbf{n} = (n_0, n_1, n_2, \ldots, n_{|M|})$ to evaluate

2. The time taken to solve a minimum cost network flow at each step.

The second, which is the computational cost incurred in solving network flow problem at each iteration of the algorithm, can be reduced with the following observation. The graph created by the algorithm is bipartite with SKUs in one partition and the modes in the other, with edges between them, In addition, in a real warehouse, the number of nodes on one partition (number of modes) is much smaller than the number of nodes in the other (number of SKUs) and therefore it is what is termed as an *unbalanced* bipartite graph. The cost scaling method for minimum cost network flows by Goldberg and Tarjan [19] as per the performance analysis by Ahuja, Orlin, Stein and Tarjan [3] for unbalanced bipartite graphs, runs in $O(|M|^2 (|N|+|M|) \log(|M|\, C)$ time, where C is the maximum value of edge cost. This is a significant improvement over the original algorithm by Orlin [29] by a factor of $|N|^2$ which tends to be very large in the case of warehouses. However, this alone is not sufficient. For example, one of the warehouses we consider in the next section is one with more than 16000 SKUs and 2 forward pick areas. The Goldberg and Tarjan algorithm takes 0.09 seconds on a fast computer for each iteration and given that there are $|N|^{|M|}$ iterations, this algorithm would take almost 250 days to complete. Even though the algorithm can be easily made to run in parallel, any such implementation on a computer with practical number of cores (usually a small number) will take many days, if not weeks. Therefore, it is impractical to run the algorithm even with the faster network flow approach, making it essential to reduce the number of iterations required to determine the optimal assignment.

The number of iterations is a first order effect on computational time and we show a way to eliminate certain assignment combinations from consideration. We do that by calculating a lower bound for the cost of the assignment combination considered in the current iteration based on what is known from prior iterations. If the lower bound is lower than the best solution found so far, we compute the exact cost of the assignment combination using the minimum cost network flow algorithm. Otherwise, we skip the combination and move on to the next iteration.

The following Proposition establishes the lower bounds for the optimal solution given the current assignment combination. For the purposes of the Proposition, we define the following notation. Let the net benefit of assigning a SKU i to mode l with n_l SKUs be defined as:

$$B(l, n_l, i) = s_l p_i - c_{rl} f_i n_l / V_l.$$

Therefore, the total net benefit for a SKU assignment is the difference between the cost of picking all SKUs from the reserve and the cost of SKU assignment under consideration, including the picking and restocking costs of SKUs assigned to modes. Therefore define optimal total net benefit $Z(\mathbf{n})$ as

$$Z(\mathbf{n}) = \sum_{i \in N} d_0 p_i - C(\mathbf{n}). \tag{5.4.1}$$

Proposition 5.4.1. Let $\mathbf{n} = (n_0, n_1, \ldots, n_{|M|})$ specify an allocation of number of SKUs assigned to each mode and $Z(\mathbf{n})$ be the corresponding optimal net benefit given $\mathbf{n}$. For a mode w, where $n_w > 0$, let $u(w)$ be the SKU that has the least net benefit among the SKUs in mode w. Let $B(w, n_w, u(w))$ be the net benefit generated by the SKU $u(w)$

in the optimal solution. For a feasible allocation $\tilde{\mathbf{n}} = (n_0 - \delta_0, n_1, \ldots, n_w + \delta, \ldots, n_{|M|})$, the corresponding optimal net benefit has an upper bound:

$$Z(\tilde{\mathbf{n}}) \leq Z(\mathbf{n}) + B(w, n_w, u(w))\delta.$$

Or equivalently,

$$C(\tilde{\mathbf{n}}) \geq C(\mathbf{n}) - B(w, n_w, u(w))\delta.$$

Proof. Let $\tau(\tilde{\mathbf{n}}, l)$ for all $l \in M \cup \{0\}$ be an assignment of SKUs to each mode l that produces the optimal net benefit, $Z(\tilde{\mathbf{n}})$. Note that $|\tau(\tilde{\mathbf{n}}, l)| = \tilde{n}_l$ for every mode l. Similarly, let $\tau(\mathbf{n}, l)$ be an assignment of SKUs in mode l that produces the optimal net benefit, $Z(\mathbf{n})$.

Note that the optimal solution for $\tilde{\mathbf{n}}$ assigns to the forward pick areas at least δ SKUs that were in the reserve in the optimal solution for $\mathbf{n}$. And mode w has at least δ SKUs in $\tau(\tilde{\mathbf{n}}, w)$ not contained in $\tau(\mathbf{n}, w)$. Therefore, there exists a set, $\Gamma \subseteq \tau(\tilde{\mathbf{n}}, w) \setminus \tau(\mathbf{n}, w)$ such that $|\Gamma| = \delta$. Consider a SKU $i \in \Gamma$ and for this SKU, we define a path through the modes, $F_i = (m_0, m_1, m_2, \ldots, m_t)$, such that the origin is $m_0 = w$ and the destination is $m_t = 0$. We suppress the dependence of m_k and t on i for notational convenience. Let SKU i be assigned to mode m_1 in the assignment $\tau(\mathbf{n}, m_1)$. If $m_1 = 0$, we are done. If not, since $i \notin \tau(\tilde{\mathbf{n}}, m_1)$ there is at least one SKU $i_1 \in \tau(\tilde{\mathbf{n}}, m_1) - \tau(\mathbf{n}, m_1)$. Let m_2 be the mode such that $i_1 \in \tau(\mathbf{n}, m_2)$. While making sure the SKUs i_1, i_2, etc. are unique to each i, repeat until eventually for some t, $m_t = 0$. Therefore for each P_i, there is a corresponding SKU path $G_i = (i_0, i_1, \ldots, i_{t-1})$ where $i_0 = i$. This has to be the case since there has to be unique sequence of moves of SKUs through the modes for each SKU $i \in \Gamma$ starting from the reserve.

Consider the optimal solution $\tau(\mathbf{n}, .)$ and to this solution, apply the set of moves for some SKU $i \in \Gamma$, (F_i, G_i), where we move SKU $i_{t-1} \in \tau(\mathbf{n}, m_t)$ to m_{t-1} and i_{t-2} to m_{t-2} and so on until SKU $i_0 = i$ is moved to $m_0 = w$, where $t = |G_i|$. We move

SKU $u(w)$ to the reserve so that mode w has the same number of SKUs as before. The objective value of this new solution has to be less than the optimal value $Z(\mathbf{n})$ and therefore net benefit due to moving SKUs has to be non-positive and so

$$B(w, n_w, i) - B(w, n_w, u(w)) + \sum_{k=1}^{t-1} B(m_k, n_{m_k}, i_k) - B(m_k, n_{m_k}, i_{k-1}) \leq 0.$$

Therefore,

$$B(w, n_w, i) + \sum_{k=1}^{t-1} B(m_k, n_{m_k}, i_k) - B(m_k, n_{m_k}, i_{k-1}) \leq B(w, n_w, u(w)). \tag{5.4.2}$$

The solution $\tau(\tilde{\mathbf{n}}, .)$ can be thought of as being derived from another assignment of SKUs $\kappa(\mathbf{n}, .)$ with net benefit K by applying the moves (F_i, G_i) for each SKU $i \in \Gamma$. The set of moves applied adds δ SKUs to mode w, removes δ SKUs from the reserve and the number of SKUs in mode $l \notin \{0, w\}$ remains the same. Therefore, we get the desired number of SKUs, $\tilde{n}$, in the resulting assignment.

Since the net benefit is decreasing in the number of SKUs and hence for any SKU i, $B(w, n_w + \delta, i) \leq B(w, n_w, i)$ we can write the optimal net benefit $Z(\tilde{\mathbf{n}})$ as

$$Z(\tilde{\mathbf{n}}) \leq K + \sum_{i \in \Gamma} B(w, n_w, i) + \sum_{k=1}^{t-1} B(m_k, n_{m_k}, i_k) - B(m_k, n_{m_k}, i_{k-1}).$$

Applying the fact that $K \leq Z(\mathbf{n})$ as $Z(\mathbf{n})$ is the optimal value and employing (5.4.2), we get

$$Z(\tilde{\mathbf{n}}) \leq Z(\mathbf{n}) + B(w, n_w, u(w))\delta.$$

The equivalent statement in terms of the cost function $C(\mathbf{n})$ follows immediately from the definition of Z in (5.4.1). $\square$

This result allows faster enumeration of all possible values for $\mathbf{n}$ in Algorithm 8. For example, let's say we are iterating through all possible assignments for $n_{|M|}$ keeping the values $n_1, \ldots, n_{|M|-1}$ constant and assigning $n_0 = |N| - \sum_{l \in M} n_l$ such that $\mathbf{n}$ is feasible. Let the cost when $n_{|M|} = 1$ be C_1 and the net benefit of the single SKU in mode $|M|$ be B. Also, let $\bar{C}$ be the lowest cost so far in the enumeration process.

From the Proposition, we see that the next value of $n_{|M|}$ to consider for evaluation is to add at least $\delta = (C_1 - \bar{C})/B$ SKUs to mode $|M|$. Now if $\delta > 1$, then we have saved the time to evaluate the minimum cost network flow for the values of $n_{|M|}$ between 2 and $1+\lceil\delta\rceil$.

In the next section we consider specific warehouses and show the effectiveness of these methods in speeding up the algorithm to determine the optimal Equal Space allocation.

5.5 A Equal Space Allocation Heuristic and Computational Performance

Algorithm 8 uses network flow for each permutation of $\mathbf{n}$ and therefore is computationally more expensive than the algorithm for Equal Space Allocation for a single forward pick area. In order to simplify the the process, we propose a heuristic which is a sequential application of the single mode Equal Space allocation Algorithm 7. The modes are sorted in descending order of pick savings. For the first mode in the list, we allocate the SKUs using Algorithm 7. Removing the SKUs allocated, we repeat it with the next mode and the remainder of SKUs until we run through all the modes. Any unallocated SKUs at the end of the heuristic is assigned to the reserve. This heuristic is described in Algorithm 9.

Note that the algorithm allocates SKUs sequentially and the modes ordered using the notion of ranking modes from the optimal labor efficient allocation. Computationally it is very efficient, as it is solving $|M|$ single mode equal space allocations. We evaluate this heuristic along with the optimal labor efficient allocation (Algorithm 4) as well as the optimal equal space allocation (Algorithm 8) for 3 warehouses described in Table 6. The warehouses each have two forward pick areas and a reserve, along with SKUs ranging from 3,049 to 16,729. The volume, pick costs and restock costs are in normalized units.

Algorithm 9: Algorithm to determine approximate assignment for a warehouse with multiple forward pick areas using Equal Space allocation.

Let the set of SKUs to be allocated be N
Sort the forward pick areas in the set M in the descending order of their pick savings, breaking ties arbitrarily. Let the forward pick areas be $M = \{1, 2, \ldots, |M|\}$.
Set $S_m = \emptyset \;\; \forall m \in M$
for $m = 1$ **to** $|M|$ **do**
    Assume mode m is the only forward mode
    Using Algorithm 7 with mode m and set of SKUs N, determine the optimal set of SKUs, S_m to be assigned to mode m
    Let $N \leftarrow N \setminus S_m$
    if $N = \emptyset$ **then** break
The set of SKUs S_m, for all $m \in M$ provides the assignment of SKUs to each forward mode m
For each SKU in mode m, the space allocation is $V_m / |S_m|$
The set of SKUs in N is the remainder to SKUs to be picked from the reserve.

The optimal labor efficient allocation (OPT) implementation is exactly as presented in Algorithm 4, enumerating all possible assignments of SKUs in modes based on their respective rankings. In each iteration, the computation is very quick, being a $O(|N|)$ evaluation of the objective function. The equal space allocation (EQS) has the same number of enumerations as the optimal solution. However, the minimum cost network flow problem is too expensive to be repeated that many times. For example, in the case of Warehouse B, each network flow problem solution time is 0.08 seconds on a Intel Core i7 processor equipped computer. We utilize the network optimization codes made available by Boris Cherkassky and Andrew Goldberg based on the methodology outlined in Goldberg [18]. The number of enumerations is of the order of 16729^2, resulting in a total solution time of almost 250 days. However, using the Proposition 5.4.1 and discussion in Section 5.4, the number of assignment combinations can be cut down to a small fraction of that number. The costs and run times of the three different methods applied to the warehouses in Table 6 is shown in Table 7.

Table 6: Parameters of three warehouses to evaluate performance of allocation heuristics. For each warehouse, we have the number of SKUs, the list of modes including the available forward pick areas and reserve. For each forward pick area, we list the the space available in each along with their respective cost per pick and cost per restock and also the cost per pick from the reserve.

Warehouse	SKUs	Storage Modes	V_m	d_m	c_{rm}
A	3,049	Flow rack	45×10^5	0.1	1.5
		Shelving	9×10^8	0.4	2
		Reserve	-	1	-
B	16,729	A-frame	1.4×10^6	0	4.75
		Knapp OSR	3.67×10^7	0.21	4.26
		Reserve	-	15	-
C	3,464	Flow rack	21,504	4	12
		Shelving	3,456	2.67	5.33
		Reserve	-	10	-

Table 7: Results of computing the near-optimal assignment for optimal allocation (OPT) using Algorithm 4, optimal assignment for Equal Space (EQS) using Algorithm 8 and the approximate assignment for Equal Space (EQSH) using the heuristic in Algorithm 9 for warehouses in Table 6. Run times are measured on a computer with Intel Core i7 quad core processor with 16GB RAM.

Warehouse	Metric	OPT	EQS	EQSH
A	Cost	63,864	71,699	74,290
	Time (s)	130	180	1
B	Cost	19,887	30,053	190,959
	Time (s)	24,161	46,187	62
C	Cost	198,097	213,524	215,853
	Time (s)	186	352	5

From Table 7, for warehouses A and C, the Equal Space allocation heuristic in Algorithm 9 produces an allocation that is within 5% of the optimal Equal Space allocation. However, in Warehouse B, we see that the heuristic is more than 600% away from the optimal Equal Space solution. This means that sometimes, the performance of the heuristic can be quite poor.

The difference between optimal solution and the optimal Equal Space allocation is similar. Warehouses A and C the total cost for the two solutions is relatively closer than for warehouse B. However, the optimal Equal Space allocation is relatively closer to optimal allocation, ranging from 8% (Warehouse C) to 52% (in Warehouse B). Also notice that though the optimal Equal Space allocation is slowest to evaluate, it is of the same order as the time required to calculate the OPT allocation. The results from Section 5.2 help speed up the Equal Space allocation so much so that the time required to solve the largest instance B is just over 12 hours and is comparable to solution times for the optimal labor efficient allocation. The algorithm for the Equal Space allocation can be easily done in parallel and this can cut down the time even further to just a few hours. The Equal Space allocation heuristic is significantly faster and takes barely a few seconds to estimate the allocation.

5.6 Equal Time allocation

The Equal Time allocation stores SKUs in the forward pick areas such their restocking frequencies are identical and can be restocked together. Bartholdi and Hackman [8] show that the Equal Time allocation model is identical to the Equal Space allocation model as represented in (5.3.1). Hence the results shown for the Equal Space apply for the Equal Time allocations. The only additional step is to determine the space allocation for each SKU i in mode m, which can be done as follows:

$$v_{mi} = \frac{x_{mi} f_i}{\sum_{k \in N} x_{mk} f_k} V_m. \tag{5.6.1}$$

5.7 Equal space allocation with additional constraints

Like in the case of the optimal labor efficient allocations, we consider constraints on number of picks and restocks that may be imposed due to warehouse layout, equipment used and workforce availability. We first address the case of single forward pick area and restate the formulation (5.2.1) with constraints on the number of restocks and number of picks as shown below in (5.7.1). The additional constraints, (5.7.1c) and (5.7.1b), ensure that the number of picks and number of restocks are limited by R and P respectively during the planning horizon.

$$\max_{\mathbf{x}} \sum_{i \in N} x_i \left(sp_i - \frac{c_r f_i \sum_{j \in N} x_j}{V_F} \right) \tag{5.7.1a}$$

subject to

$$\frac{\left(\sum_{i \in N} x_i f_i\right)\left(\sum_{j \in N} x_j\right)}{V_F} \leq R \tag{5.7.1b}$$

$$\sum_{i \in N} x_i p_i \leq P \tag{5.7.1c}$$

$$x_i \in \{0, 1\}, \quad i \in N \tag{5.7.1d}$$

Using the notation $n_F = \sum_{j \in N} x_j$ from Section 5.2, we rewrite the problem as follows:

$$\max_{\mathbf{x}} \sum_{i \in N} x_i \left(sp_i - \frac{c_r f_i n_F}{V_F} \right) \tag{5.7.2a}$$

subject to

$$c_r \frac{\left(\sum_{i \in N} x_i f_i\right) n_F}{V_F} \leq R \tag{5.7.2b}$$

$$\sum_{i \in N} x_i p_i \leq P \tag{5.7.2c}$$

$$x_i \in \{0, 1\} \qquad\qquad i \in N \tag{5.7.2d}$$

For a fixed n_F, this is a two dimensional knapsack problem which admits a Polynomial Time Approximation Scheme (PTAS) [16]. However, no Fully Polynomial

Time Approximation Scheme (FPTAS) exists unless P=NP [24] for k-dimensional knapsack problems, where $k \geq 2$. This implies that the ϵ-approximation schemes are inefficient, but still are practical when the number of SKUs is large. However, in the special case where just one of the two knapsack type constraints is present, the problem reduces to a single dimensional knapsack problem which admits FPTAS. In addition, the simple greedy algorithm for the knapsack problem determines a solution that is within net benefit of one SKU from optimal. When n_F is a large number, as is usually the case in most warehouses, this provides a very good solution [5].

However, for the general case, the news is not all bad. Multidimensional knapsack problems appear in many settings and has attracted the attention of researchers for a long time. Freville [15] provides a survey of the main results published in the literature on the topic. When the knapsack dimension is fixed, as is in our case, many approximation algorithms exist that can produce effective solutions, with some of them shown to be probabilistically near-optimal.

The algorithm to determine an optimal assignment with Equal Space allocation in the case of a warehouse with one forward pick area with constraints on number of picks and restocks is as follows. For each $n_F \in 1, \ldots, |N|$, we solve the problem as formulated in (5.7.2) by using an algorithm as applicable based on the dimensionality of the resulting knapsack problem. Choose the value of n and the corresponding solution that generates the maximum objective value. The complexity of this method is $|N|$ times the computational time required by the algorithm to solve the knapsack problem for a given n_F.

We also consider the extension of warehouses with multiple forward pick areas that have constraints on the number of picks and restocks that can be achieved within the planning horizon. The formulation to determine the optimal assignment for such a warehouse using Equal Space allocation is shown in (5.7.3). Each forward pick area has a constraint on the number of restocks (5.7.3c) and the number of picks

(5.7.3e) that can be accomplished during the planning horizon. In addition, there are constraints on the total number of picks (5.7.3f) and total number of restocks (5.7.3d) across all forward pick areas.

$$\max_{\mathbf{x}} \sum_{i\in N, l\in M} x_{li}\left(s_l p_i - \frac{c_{rl} f_i \sum_{j\in N} x_{lj}}{V_l}\right) \tag{5.7.3a}$$

subject to

$$\sum_{l\in M} x_{li} \leq 1 \qquad i \in N \tag{5.7.3b}$$

$$c_{rm}\frac{\left(\sum_{i\in N} x_{mi} f_i\right)\left(\sum_{j\in N} x_{lj}\right)}{V_m} \leq R_m \qquad \forall m \in M \tag{5.7.3c}$$

$$\sum_{l\in M} c_{rl}\frac{\left(\sum_{i\in N} x_{li} f_i\right)\left(\sum_{j\in N} x_{lj}\right)}{V_l} \leq R \tag{5.7.3d}$$

$$\sum_{i\in N} x_{mi} p_i \leq P_m \qquad \forall m \in M \tag{5.7.3e}$$

$$\sum_{\substack{i\in N\\ l\in M}} x_{li} p_i \leq P \tag{5.7.3f}$$

$$\sum_{i\in N} x_{mi} f_i \leq F_m \qquad \forall m \in M \tag{5.7.3g}$$

$$x_{mi} \in \{0,1\} \qquad i \in N, m \qquad \in M \tag{5.7.3h}$$

For a fixed $n_l = \sum_{j\in N} x_{lj}$ for all modes $l \in M$, this formulation represents the Multi-Resource Generalized Assignment Problem (MRGAP) which is strongly NP-hard and does not admit a PTAS [24]. The additional knapsack-type constraints make it a multi-dimensional multi-knapsack problem that at best admits only pseudo-polynomial algorithms based on dynamic programming [24]. The complexity of these algorithms can be very large when $|N|$ is large, which is usually the case in our setting. Only in the case when all forward pick areas are completely identical in picking and restocking costs and we have only one of the three types of constraints with the limit also being exactly the same for all modes, do we have a PTAS. In any other case, even with just two distinct types of forward pick areas (either the pick cost or the restock cost or the capacity is different) with only one of the two types of constraints, no PTAS exists unless P=NP [10]. However, there are constant factor approximation algorithms in the case where we have exactly one of two constraints (either on the number of picks or restocks), with the best known to be a 2-approximation algorithm [10]. Assuming an algorithm is available to solve the MRGAP for every feasible $(n_l)_{l \in M}$, the methodology would proceed exactly as before for the single forward pick area case.

As noted earlier, Equal Time allocations have the same objective as Equal Space, and so the analysis remains the same. We see that in the case of Equal Space and Equal Time allocations, the additional constraints make it significantly harder to find an optimal or near-optimal solution.

5.8 Conclusions

In the case of Equal Time and Equal Space allocations without additional constraints on picks, flow or restocks, we present algorithms to obtain the optimal assignment of SKUs both for single and multiple forward pick area cases. This is useful in cases where warehouses are constrained by their space allocation strategy but are considering which SKUs to store in the forward area. The Equal Space allocation allows for easy space management in the forward area and the Equal Time allocation allows for simultaneous restocks if the entire forward pick area has to be restocked at once. However, none of the algorithms show an a priori ranking of SKUs unlike in the case of the optimal allocations nor unimodality properties. This means that all possibilities in terms of number of SKUs in each forward pick area have to be evaluated. Furthermore, in case of any changes to the parameters (SKUs, picks, flow or forward pick area attributes), the new optimal assignment determined can be very different from the previous one.

We also provide a heuristic solution for the Equal Space allocation in warehouses with multiple forward pick areas. With the addition of constraints on picks, flows or restocks, we frame the equal space (or time) allocation in a warehouse with a single forward pick area as a multi-dimensional knapsack problem, which allows a polynomial time approximation scheme in addition to a pseudo-polynomial algorithms. Though these are generally more difficult to solve, a variety of heuristics are available [15] that can be used to get solutions of good quality. In the case of the equal space (or time) allocation in multiple forward pick areas, the resulting multi-dimensional knapsack problem is much harder to solve.

We also test the algorithms for labor efficient allocation, optimal equal space allocation and heuristic equal space allocation on data sets from 3 warehouses with varying number of SKUs and modes. In each case, the methodology we present runs in reasonable time, even for large warehouses. The heuristic equal space allocation though runs in quick time, the solution quality as compared to the optimal varies in each of the 3 cases considered.

Like in the case of the the algorithms to determine the optimal allocation, the Equal Time and Equal Space allocation algorithms allow for many generalizations. They can be easily extended to allow for pick savings or restock costs to vary by SKU or forward pick area. This allows for flexibility in the cost model that may be dependent on the nature of the SKU, the type of forward pick area (flow rack vs. rack shelving) or the location within the warehouse that could have higher travel times. Furthermore, the algorithms also accommodate the possibility that some SKUs can be stored only in a certain set of forward modes or may have a minimum space (reorder points) that need to be allocated depending on the nature of the SKU and the respective forward mode. Lastly, since the mathematical formulations for Equal Space and Equal Time are identical [8], the algorithms also apply to warehouses that use a combination of the two allocation strategies, i.e., certain forward pick areas are allocated on the basis of Equal Space and the rest on the basis of Equal Time.

CHAPTER VI

POWER LAWS IN WAREHOUSES

6.1 Introduction

Power laws have been observed in many natural and man-made phenomena. One of the earliest observations of power law behavior was by Kingsley Zipf, a professor of linguistics, who sought to determine the frequency of use of words in an English text. He observed that the frequency of a word is inversely proportional to its rank when sorted in descending order of frequency, giving rise to a special case of the power law called Zipf's law. Such power laws have subsequently been observed in areas as diverse as human population in cities, income distribution, internet traffic, natural languages, corporation sizes etc.

Based on Zipf's work, the Power Law for a quantity is described as follows. If r is the rank of quantity Q in descending order, Q is proportional to inverse of r raised to the power γ, where $\gamma > 0$. Expressed in mathematical form,

$$Q \propto r^{-\gamma}. \tag{6.1.1}$$

Zipf's observation on the distribution of frequency of words corresponds to a Power Law with $\gamma = 1$. This special case is named Zipf's law, after him.

An alternative form is the power law distribution, $p(x)$, that is defined as follows:

$$p(x) \propto x^{-\beta}, \tag{6.1.2}$$

where β is called the *scaling parameter* of the power law. Adamic [1] showed that the two forms, (6.1.1) and (6.1.2), are mathematically equivalent and the relationship between β and γ is

$$\beta = 1 + \frac{1}{\gamma}.$$

He also shows the Pareto distribution is equivalent to the power law by definition. The Pareto distribution is usually defined as

$$P[X > x] \propto x^{-\beta+1},$$

which is same as the complementary cumulative distribution function of the power law in (6.1.2).

In the case of warehouses the ABC distribution is frequently invoked to describe the popularity distribution of SKUs. In fact, inventory is frequently managed based on the assumption of such a distribution. The ABC distribution can be viewed as an approximation of the Pareto distribution and therefore, it is important to understand if the distribution of SKU movements is actually a power law.

In this chapter, using data from different warehouses, we analyze whether SKU popularity, i.e., the number of picks by SKU, fits a power law distribution. Though there have been studies of the not-so-popular SKUs and their contribution to total business of warehouses [4], in practice the SKUs with little demand are usually discontinued thereby affecting the distribution. Therefore, we focus on the upper tail of the distribution comprising of the most popular SKUs where we identify the best power law fit, estimate statistical support for the power law and test whether the power law could be a more plausible fit than other alternative distributions. We analyze the warehouses by industry category to determine if any similarities in the fitted power law exist. We also review power law generation mechanisms that could serve as plausible explanations for emergence of power laws in warehouses.

6.2 About Power Laws

The power law probability density function for a continuous random variable X given in (6.1.2) is as follows:

$$p(x) \propto x^{-\gamma-1} = Cx^{-\beta}. \tag{6.2.1}$$

This function diverges as $x \to 0$ and therefore, the distribution can hold only for $x > x_{\min}$ where $x_{\min} > 0$ is a lower bound on the distribution. There the normalizing constant C is

$$C = (\beta - 1)x_{\min}^{\beta-1}.$$

The corresponding Pareto distribution is as follows:

$$\Pr[X > x] = (\frac{x}{x_{\min}})^{-\beta+1}.$$

Note that this distribution produces a straight line when plotted in a log-log scale.

In the event X is a discrete random variable, the power law distribution is defined as

$$\Pr[X = x] \propto x^{-\gamma-1} = Cx^{-\beta}. \tag{6.2.2}$$

As before, C is a constant, $\beta > 1$ since $\gamma > 0$ and the density diverges at $x = 0$. Therefore, the power law can hold only for $x > x_{\min}$. Calculating the normalizing constant, C,

$$C = 1/\zeta(\beta, x_{\min}),$$

where,

$$\zeta(\beta, x_{\min}) = \sum_{n=0}^{\infty}(n + x_{\min})^{-\beta}$$

is the generalized or Hurwitz zeta function [11]. The corresponding Pareto distribution is as follows:

$$\Pr[X > x] = \frac{\zeta(\beta, x)}{\zeta(\beta, x_{\min})}.$$

Power law distributions have two main characteristics. One is that they are heavy tailed distributions, where the extreme events have a significantly higher probability of occurrence compared to distributions like normal or exponential. The second is that they are *scale-free*, in that, the distribution is the same even when a power law distributed variable is scaled by a constant. Newman [28] defines it in mathematical terms as follows:

$$p(bx) = g(b)p(x),$$

where, b is a constant and $g(.)$ is a function. Therefore even if we measure x in a different unit, the power law distribution shape is unchanged except for a multiplicative constant. In fact the power law distribution is the only scale-free distribution.

6.3 Power Law with exponential cut-off

The power law distribution with exponential cut-off is a generalization of the power law. The probability density function is as follows.

$$p(x) \propto x^{-\beta} e^{-\lambda x}.$$

where $\beta > 1$ and $\lambda > 0$ are the parameters of the distribution.

We observe that that when $\lambda = 0$ this expression becomes the power law. When $\lambda > 0$, the exponential term diminishes the tail and this distribution is not a real power law. However, in the regime of $x \ll 1/\lambda$, the distribution behaves like a power law. When fitting power laws to data sets, we check whether the power law is a more plausible fit than the power law with exponential cut-off. Other alternative distributions considered are the lognormal, exponential, and stretched exponential distributions. In the case of discrete data sets, we consider the Poisson distribution as well.

6.4 Is it a Power Law?

We analyze picks-per-SKU data from 30 warehouses for power law behavior in the upper tail of the distribution. These warehouses vary across the spectrum in terms of goods shipped (sunglasses, groceries, office supplies, auto spare parts among others), small and large SKUs, and types of shelving, layout and level of automation employed. We use the methodology described in Clauset et al. [11] to fit a power law distribution in the upper tail and to test the quality of fit. The methodology is as follows:

1. *Fit Power Law*: We estimate the parameters of the power law distribution that fits the data using maximum likelihood estimation (MLE) methods. Depending

on the data, we fit the discrete (if the data is integral and smaller than 1000 data points) or the continuous (all other cases) form of the power law distribution. We calculate the parameter $x_{\min}$, which is the cut-off above which the power law distribution is fitted, and β, the scaling parameter of the distribution.

2. *Goodness of fit*: Estimate the p-value of the power law distribution as defined by Clauset et al. [11] as follows. We use the power law distribution parameters estimated in step 1 to generate many synthetic data sets, to which we in turn fit a power law distribution. We calculate a Kolmogorov-Smirnov goodness of fit statistic for each fit. The p-value is the fraction of synthetic data sets with the goodness of fit statistic less than that of the original data set. This is interpreted as the probability that the power law distribution estimated in step 1 fits the sample. Following the recommendation of Clauset et al. [11], we set a significance level of 0.1 and rule out power law fits for distributions with p-values less than 0.1.

3. *Test alternative distributions*: We also fit other distributions to the data and compare the fit to the power laws using likelihood ratios. If the log-likelihood ratio is positive with the p-value less than 0.1, we rule out the null hypothesis and take it that the power law is a more plausible fit than the competing distribution. If the log-likelihood ratio is negative with a p-value less than 0.1, then the alternative distribution is a more plausible fit than the power law.

Using the methodology of Clauset et al. [11], we test for the evidence for power laws in the upper tail on picks-per-SKU data from 30 different warehouses. They are listed below with a label and description.

Table 8: Description of 30 warehouses and the type of items stocked and shipped from each. Each is assigned a name in the first column and the second describes the warehouse and nature of items stocked.

Warehouse ID	Warehouse Type
Apparel 1	Ships low-end apparel to large retailers
Apparel 2	T-shirts
Apparel 3	Ships t-shirts to small retailers
Cold chain	Cold chain distribution center
Dairy products	Dairy products
Grocery 1	Supports a grocery chain
Grocery 2	Supports another grocery chain
Grocery 3, toxic	Fulfills e-commerce orders for non-food grocery items
Hardware catalog 1	Hardware distributor
Hardware catalog 2	Second hardware distributor
Hardware catalog 3	Third hardware distributor
Industrial controls	Industrial controls
Personal care 1	Personal care items
Personal care 2	Personal care items
Personal care 3	Fulfills direct customer orders for personal care items
Pharma 1	Supports a pharmacy chain
Pharma 2	Supports another pharmacy chain
Pharma 3	Supports a third pharmacy chain
School biology labs	Biology lab kits for schools
School science labs	Science lab kits for schools
Service parts 1, auto	Automotive service parts
Service parts 2, trucks	Service parts for trucks and trailers
Service parts 3, trucks	Service parts for heavy machinery and trucks
Service parts 4, trucks	Service parts for a parcel delivery truck fleet
Soda 1	Beverages, bottles and cans
Soda 2, small orders	Beverages, carton orders only
Soda 2, unit loads	Beverages, unit load orders only
Supply 1, office	Office supplies
Supply 2, school	School supplies
Wine	Wine cases and bottles

These data sets represent number of picks by SKU for a time period at least one month, so that the underlying distribution is stable. In Table 9 we show the power law fits to each of the data sets listed above, along with their basic statistics: the total number of data points n, mean $\bar{x}$, standard deviation σ, and maximum value $x_{\max}$ of the data set. We include the power law fit with the estimates of the two parameters, $x_{\min}$ and β, along with the corresponding standard errors based on the methodology presented in Clauset et al. [11]. In addition, we also tabulate the value n_{tail} (and its standard error), which is the number of data points above $x_{\min}$ (the upper tail where the power law is fitted).

Each of the data sets is classified as *discrete* or *continuous* so that the corresponding form of the power law distribution is fitted. We follow Clauset et al. [11] and classify those data sets that are integral as discrete and the rest as continuous. However, data sets that entirely consist of integers are classified as continuous if they have a minimum value of 1000 and at least 100 data points.

The number in brackets alongside the values for β in Table 9 is the uncertainty in the final digit. For example, the estimated fit for Apparel 1 warehouse is a power law distribution with $x_{\min} = 1993 \pm 504$ and $\beta = 2.18 \pm 0.09$. The number of points above $x_{\min}$, or n_{tail} is 330 ± 119. The last column is the p-value indicating the validity of the power law fit. We consider any value over 0.1 to indicate a good power law fit to the data set and mark it in bold in the table. Out of the 30 data sets, 16 have a p-value of 0.1 or higher. For these 16 data sets, the power law is a plausible fit.

From the power law fits, we observe that the power law scaling factor β has a wide range of possible values depending on the data set, the lowest being 2.0 for the Cold Chain, Soda 1 and Soda 2 warehouses and the highest 5.2 for Pharma 3. Note that these values have some uncertainty in the final digit and therefore that the range could be even broader. Note that the size of the upper tail region where the power law is fit as indicated by $x_{\min}$ and n_{tail} is much smaller than the whole data set in almost every instance. This is often the case in real life data sets like warehouses for example, where SKUs with low picks maybe periodically eliminated and new SKUs introduced, possibly affecting the distribution.

Table 9: Description of picks-per-SKU data from 30 warehouses along with their power law fits. Included in order of columns are the name of the warehouse, total number of SKUs, statistics on the number of picks by SKU (mean, variance and maximum value) and the power law fit ($x_{\min}$, β, and number in upper tail, n_{tail}) with the corresponding p-value. Statistically significant p-values are marked in **bold**.

Warehouse	n	$\bar{x}$	σ	$x_{\max}$	$x_{\min}$	β	n_{tail}	p
Apparel 1, big retail	2189	1357.5	3804.6	39904.0	1993.0±504.0	2.18(9)	330±119	**0.26**
Apparel 2, t-shirts	8716	75.8	217.9	6511.0	93.0±18.2	2.32(4)	1584±319	**0.16**
Apparel 3, t-shirts	1000	38.0	62.3	756.0	47.0±15.5	2.5(2)	216±78	**0.28**
Cold chain	174	10.2	11.8	68.0	5.0±3.8	2.0(4)	105±25	0.00
Dairy products	451	144.1	152.2	888.0	282.0±34.2	3.5(2)	71±27	0.04
Grocery 1	89	21.5	12.6	72.3	15.0±4.2	3.0(7)	57±15	**0.10**
Grocery 2	2211	14.3	6.5	112.4	16.8±2.4	4.7(2)	441±272	0.03
Grocery 3, toxic	267	12.1	3.6	23.9	5.8±1.0	2.3(9)	254±36	0.00
Hardware catalog 1	12386	2.2	5.8	199.0	8.0±1.2	2.77(8)	697±250	**0.84**
Hardware catalog 2	14803	26.4	100.4	3442.0	149.0±46.0	2.4(2)	488±951	**0.20**
Hardware catalog 3	39698	1.5	1.3	41.0	4.0±0.7	3.50(9)	1912±2521	**0.14**
Industrial controls	5511	12.0	28.6	616.0	23.0±16.5	2.4(2)	713±254	0.00
Personal care 1	935	302.3	621.9	7868.0	789.0±216.2	2.8(4)	100±122	**0.54**
Personal care 2	875	175.0	280.5	2103.0	724.0±178.2	3.5(5)	51±84	**0.41**
Personal care 3	4746	184.9	701.0	16004.0	667.0±169.9	2.5(1)	288±115	**0.86**
Pharma 1	6231	148.4	97.6	806.0	169.0±2.5	3.50(0)	2422±67	0.00
Pharma 2	3172	167.4	410.9	9083.7	228.3±26.9	2.30(6)	568±54	**0.38**
Pharma 3	4127	2343.2	1609.9	14402.0	5700.0±1485.1	5.2(9)	215±509	**0.12**
School biology labs	7437	15.5	36.0	1516.0	30.0±10.9	2.6(1)	854±434	**0.10**
School science labs	6976	28.6	51.0	1082.0	34.0±19.1	2.5(2)	1602±371	0.00
Service parts 1, Auto	59608	6.6	12.2	219.0	46.0±13.6	3.5(5)	1277±10983	0.00
Service parts 2, trucks	16654	39.2	101.8	2082.0	217.0±89.0	2.9(3)	723±274	0.00

Table 9: Continued on next page

Table 9: Continued from previous page

Warehouse	n	$\bar{x}$	σ	$x_{\max}$	$x_{\min}$	β	n_{tail}	p
Service parts 3, trucks	80301	5.3	18.2	878.0	5.0±4.0	2.11(6)	17040±4491	0.00
Service parts 4, trucks	36101	78.6	450.7	19490.0	1966.0±710.6	2.8(4)	247±1688	**0.95**
Soda 1	759	4111.0	8872.7	100722.0	2972.0±4255.7	2.0(4)	239±72	0.00
Soda 2, small orders	364	302.0	387.3	1587.0	722.0±195.4	3.5(5)	61±37	0.01
Soda 2, unit loads	338	368.1	444.6	1654.0	217.0±182.9	2.0(5)	155±44	0.00
Supplies 1, office	20730	36.5	101.1	3537.0	143.0±30.6	2.7(1)	1025±425	**0.18**
Supplies 2, school	15357	156.1	424.2	12395.0	1791.0±808.4	3.5(7)	198±1852	**0.93**
Wine	158	46.6	57.9	357.0	33.0±25.6	2.2(5)	68±22	0.00

To analyze the warehouses by their power law fits, we group the warehouses by industry as the types of SKUs stocked and their economics are similar within each industry category. We expect that there to be similarities in whether power laws fit the category and potentially in the estimated parameters of the fit as well. The groupings and the corresponding observations on the power law fits by group are as follows:

Apparel This category refers to the Apparel 1, 2 and 3 warehouses. These data sets produce a power law fit with similar scaling factors of 2.18 ± 0.09, 2.32 ± 0.04 and 2.5 ± 0.2 respectively. Importantly, all three of the power laws are significant as shown by the p-value of greater than 0.1. Note that the order of magnitude of values in each data set is different as seen from the respective mean values ($\bar{x}$) and maximum values ($x_{\max}$). The number of data points in the upper tail is also significant and approximately in the same proportion, at about 20% of the total number of observations.

Cold chain, dairy and grocery We include in this set, the Cold chain, Dairy and Grocery 1, 2 and 3 warehouses. Note that these warehouses with the exception of Grocery 2 have the least number of observations (SKUs) among the 30 data sets. Also note that the power law scaling factor varies over a broad range of 2.0 to 4.7. The p-value, or the probability that the power law fits the sample, is below the 0.1 significance level for the warehouses in this category, with the exception of Grocery 1, which has a p-value of 0.1. However, the number of observations in Grocery 1 is only 89 lower than the minimum of 100 points that Clauset et al. [11] recommend for the statistical tests to avoid false positives. Based on the evidence at hand, the support for power laws among these data sets is quite weak.

Hardware catalog This group includes the Hardware catalog 1, 2, and 3 data sets.

These warehouses have over 10,000 SKUs as one might expect in a hardware catalog order fulfillment warehouse. The power law scaling parameters for the 3 warehouses are 2.77 ± 0.08, 2.4 ± 0.2 and 3.50 ± 0.09, indicating a broader range of possible values compared to Apparel. In all three cases, there is significant support for a power law fit as indicated by p-values of at least 0.14.

Personal care This group includes the Personal care 1, 2 and 3 warehouses. These data sets fit a power law with scaling parameter 2.8 ± 0.4, 3.5 ± 0.5 and 2.5 ± 0.1, showing a broad range. More importantly, there is significant support for the power law fit with a minimum p-value of 0.41.

Pharma This group includes the Pharma 1, 2 and 3 warehouses, with several thousand SKUs each. Again the scaling parameter for the power law fits show a wide range: 3.5, 2.30 ± 0.06 and 5.2 ± 0.9. Pharma 2 and 3 data sets also exhibit significant statistical support for the power law fit, but Pharma 1 does not. These warehouses supply to different retail pharmacy chains that stock an ever wider range of SKUs beyond pharmaceutical goods. This could be the reason for the wide range in the estimates of the power law parameters fitted.

School lab kits We include in this category the following two data sets, School biology labs and School science labs. The scaling parameters for the two warehouses are very similar at 2.6 ± 0.1, and 2.5 ± 0.2 respectively. The two warehouses that supply kits to school labs appear to have similar scaling parameters. Interestingly, only the School biology labs data set shows statistical support for the power law fit.

Service parts This category includes Service parts 1, 2, 3 and 4 warehouses. The scaling parameters for the power law fits shows a wide range: 3.5 ± 0.5,2.9 ± 0.3,2.11 ± 0.06 and 2.8 ± 0.4. Only Service parts 4 shows significant statistical support for the power law fit.

Soda This category includes 3 data sets from two warehouses Soda 1 and Soda 2. There are two data sets from the Soda 2 warehouse, one for small orders and the other for unit load orders. The Soda 2, small orders data set has an estimated scaling parameter that is much higher than the other two data sets and this could possibly be because of the fewer than 100 data points on the upper tail (n_{tail}) that can cause over estimation of the power law parameters [11]. None of the data sets in this category show any support for the power law fit.

Supplies This category includes the two data sets, Supplies 1 (office supplies) and 2 (school supplies). We have classified them together as they stock large number of similar items that are ordered by customers that are large organizations. The office supplies stocks stationery, whiteboards, and furniture among other items, whereas the school supplies warehouse stocks generic school items like writing boards, stationary, protractors etc. However, the scaling parameter for their respective power law fits differs significantly between the two warehouses, at 2.7 ± 0.1 for Supply 1 and 3.5 ± 0.5 for Supply 2. Both data sets show significant statistical support for the estimated power law fit.

Others The data sets that are not included in the above categories are: Industrial controls and Wine. Neither data set shows a significant support for the corresponding power law fit.

Broadly, we observe that in Apparel, Hardware catalog, Personal care, Pharma, and Supplies categories show statistical support for a power law fit and the other categories do not. This indicates that there could be an industry dependent effect on the emergence of power laws. On the other hand, statistical significance of power law fits does not appear to depend much on number of SKUs in the data set. For example, data sets that have statistical support for the power law with p-value is greater than 0.1, the size of the data set varies from 875 (ignoring Grocery 1 as it has fewer than 100 data points) to almost $40,000$. None of the small data sets (fewer than 500 SKUs) fit a power law and there are a few exceptions on the large data sets as well. For example, Service parts 1 and 3 have almost $60,000$ and $80,000$ SKUs but show no support for the power law fit. However, we are unable to draw any similarities of scaling parameters within a category. Only the Apparel category appears to have a moderate range of values for β, but the other categories seem to have very wide ranges.

In Figure 12 and Figure 13, we show the empirical complementary cumulative distribution function and also the estimated power law fit for each data set. The lower tail of the distribution typically is outside the power law regime causing the flat curves in the plots in all the data sets. This is possibly because warehouses discontinue SKUs with low demand and introducing new ones whose demand is still evolving. The power law applies only above the threshold that we estimate as $x_{\min}$ from where the straight line indicating the power law fit is shown. Note that the power law fit estimated using maximum likelihood methods as shown in Clauset [11] is not always what one might interpret from looking at the graphs. Even though the plots for a few data sets like Grocery 2 and 3 appear to be straight lines, we see that there is no statistical support for the power law. Thus visual interpretation alone is insufficient when fitting power laws.

In several data sets, the empirical distribution falls off at high values of x as seen in the Service parts 1, 2, and 3 as well as the Hardware Catalog 1, 2, and 3 warehouses. This is attributed to what is described as *finite-size* effects, meaning that the demand for the high sales products can only go up to a certain point after which market size and other external factors act as a constraint. This prevents the number of picks in the upper tail of the distribution from growing to the numbers as predicted by the power law. In such cases, the power law distribution with exponential cut-off could potentially fit the data much better. In the next set of results, we evaluate whether alternative probability distributions including the power law with exponential cut-off are a more plausible fit than the power law.

6.5 Do other distributions fit the data better?

Though power laws can show up as a good fit to the data set, it is entirely possible that an alternative distribution can fit the data as well or better. We test whether another distribution can be more suited to the data set in this section. The alternative distributions we consider and the statistical tests to compare against the power law for each depend on whether the data set is classified as *discrete* or *continuous*. Recall that we classify as discrete those data sets that take only integral values and the rest as continuous. As we are looking at the number of picks-per-SKU in each warehouse, being integer, we expect the distribution to be naturally classified as discrete. However, as seen in Table 11, 5 out of 30 data sets are classified as continuous. In 4 of them, the observed statistic is the mean picks-per-SKU and hence fractional numbers appear. The fifth, Pharma 3 though is a purely integral values, has over a 100 data points with the smallest value being over 1000. This we classify as continuous as the numbers are large enough to be scaled. The remaining 25 data sets are classified as discrete and results tabulated in Table 10.

To compare whether the alternative distribution could be a better fit for the data, we compare the likelihood of the data under the two distributions, the power law and the alternative. The distribution with the higher likelihood is the more plausible fit to the data. We perform the comparison is done by computing the log-likelihood ratio (LR) or the logarithm of the ratio of the two likelihoods. If this value is positive, then the power law is a more plausible fit and if it is negative, the alternative distribution is a better fit than the power law. As per Clauset et al. [11], we also compute a p-value for the LR and consider the value of the ratio as statistically stable only if the p-value is less than the significance level of 0.1. The cases where an alternative

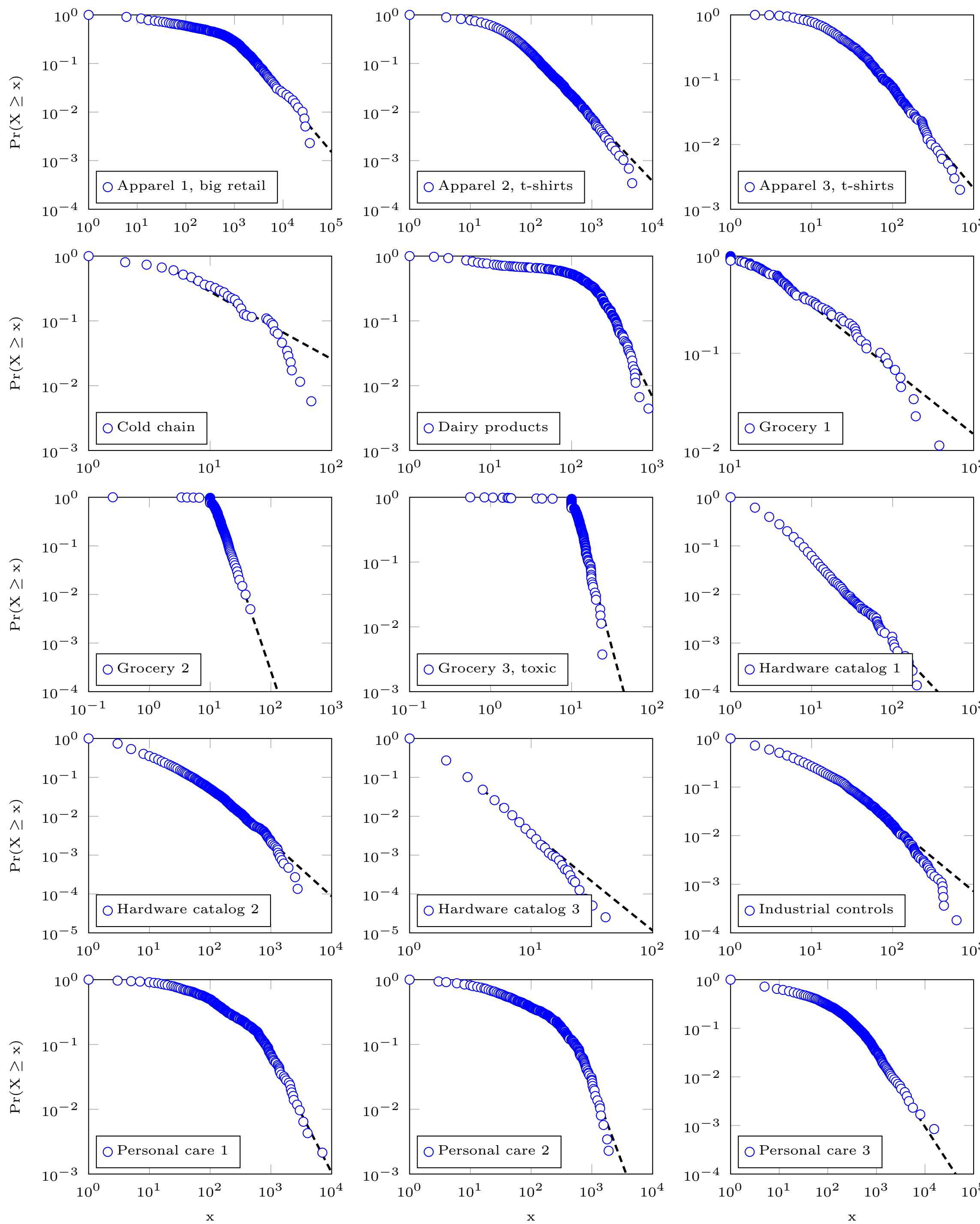

Figure 12: The empirical complementary cumulative distribution function (blue) and their corresponding estimated power law fits (black) for the first 15 of the 30 data sets.

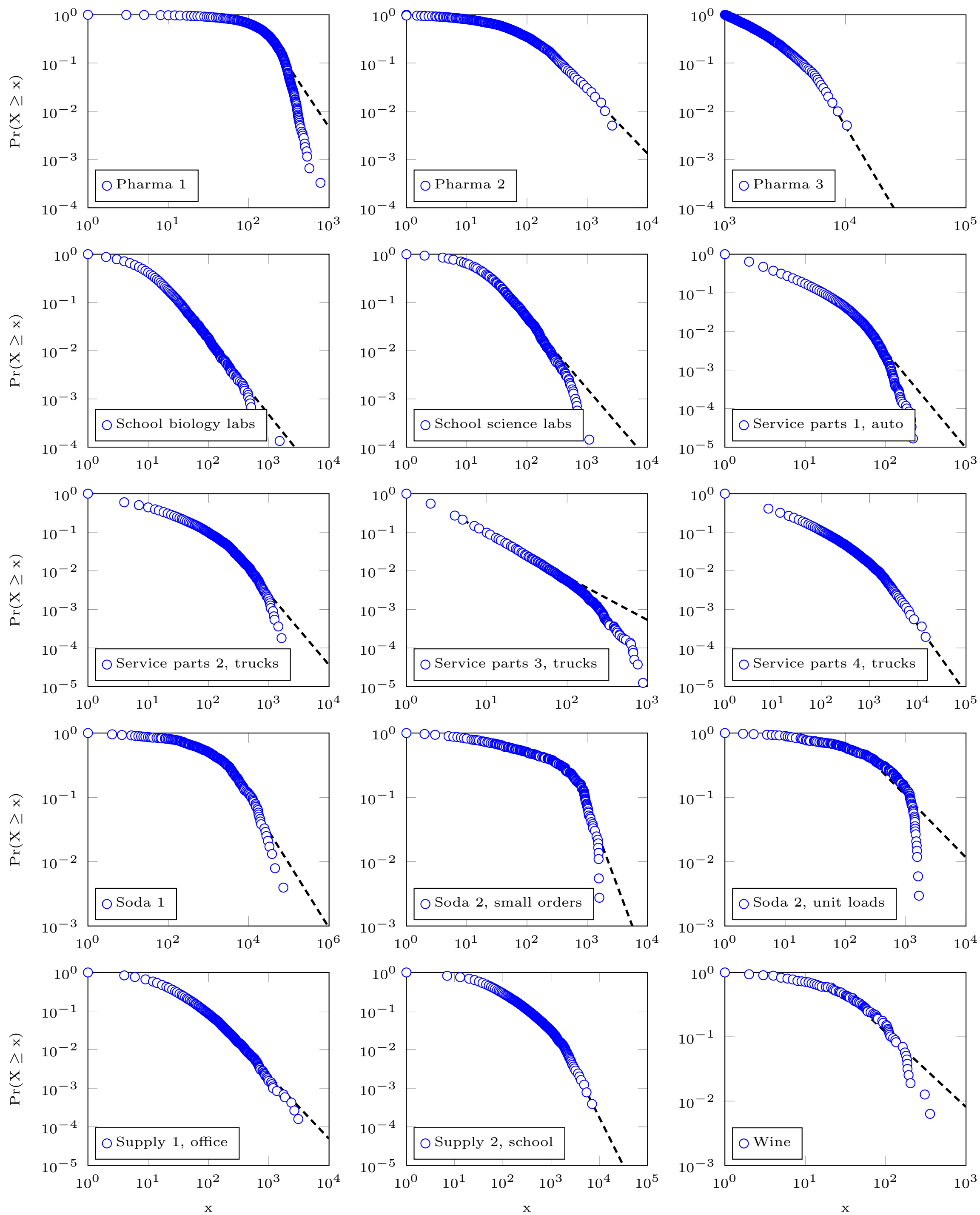

Figure 13: The empirical complementary cumulative distribution function (blue) and their corresponding estimated power law fits (black) for the second 15 of the 30 data sets.

distribution has a negative LR value and p-value less than 0.1 are marked in **bold** in Table 10 and Table 11.

Table 10 and Table 11 show the LR for the alternative distributions as compared to the power law, alongside the p-value of power law fits. We also make our observation for each data set in the column labeled Evidence. *Good* implies that the power law is the most plausible fit among the distributions considered; *Moderate* implies that though the power law is a plausible fit, other plausible alternatives exist; *Alternatives* implies that alternative distributions possibly fit the data better than the power law, and *None* implies that none of the distributions considered are plausible. We evaluate lognormal, exponential, stretched exponential and power law with exponential cut-off distributions as alternative candidate distributions for the data sets classified as continuous. For the discrete data sets, we also consider the Poisson distribution.

Overall, out of 30 data sets, only 4 data sets are best fitted by a power law and we note them as *Good* in Table 10. 12 data sets are marked *Moderate*, and in all of them, it is the power law with exponential cut-off that is seen as a more plausible distribution. Among the 13 data sets labeled *Alternatives*, the power law with exponential cut-off is the one common distribution that is more plausible than the power law. Together, this means that in 25 out of the 30 data sets, the power law with exponential cut-off is seen as a more plausible distribution than the power law. This is not surprising, considering that this distribution is a generalization of the power law. This causes the power law with exponential cut-off to produce a higher likelihood value than the power law, at least in cases where the power law fit is not plausible. The second reason is the general behavior in the data sets of the distribution diminishing faster than a power law for high values of x as seen in their graphs.

The second most common distribution that has some statistical support among the data sets is the lognormal distribution. In one data set marked *Moderate*, Hardware catalog 3, the lognormal distribution in addition to the power law with exponential cut-off is a more plausible fit than the power law. Among the 13 data sets labeled as *Alternatives*, 12 show more statistical support for the lognormal distribution than the power law. This however, does not mean that the lognormal distribution is the best fit for the data set.

The exponential and the stretched exponential distributions also appear in a few cases to indicate a better fit than the power law. The only distribution that does not pass the statistical tests for any data set is the Poisson distribution.

Table 10: Tests of power law behavior and plausibility of alternative distributions in discrete data sets. For each data set, we give a p-value for the Power Law (PL) fit and the Log-likelihood Ratio (LR) for the alternative distributions considered: Lognormal, Exponential (Exp.), Stretched Exponential (Str. Exp.), Poisson and Power Law with Exponential Cut-off (PL+cut-off) distributions. We also include the p-values that indicate the significance of each of the likelihood ratio tests. Statistically significant p-values are in **bold**.

Context	PL	Lognormal		Exp.		Str. Exp.		Poisson		PL+cut-off		Evidence
	p	LR	p	LR	p	LR	p	LR	p	LR	p	
Apparel 1, big retail	**0.26**	-1.46	0.15	5.66	0.00	-1.51	0.13	9.01	0.00	-3.89	**0.01**	Moderate
Apparel 2, t-shirts	**0.16**	-1.38	0.17	7.83	0.00	-0.23	0.81	8.12	0.00	-4.30	**0.00**	Moderate
Apparel 3, t-shirts	**0.24**	-1.16	0.25	2.37	0.02	-1.19	0.24	4.82	0.00	-2.12	**0.04**	Moderate
Cold chain	0.00	-2.57	**0.01**	-1.37	**0.17**	-2.73	**0.01**	4.76	0.00	-9.62	**0.00**	Alternatives
Dairy products	0.04	-1.61	0.11	-2.03	**0.04**	-1.82	**0.07**	4.07	0.00	-3.40	**0.01**	Alternatives
Hardware catalog 1	**0.84**	-0.69	0.49	5.60	0.00	4.19	0.00	5.60	0.00	-0.18	0.55	Good
Hardware catalog 2	**0.20**	-1.58	0.11	4.34	0.00	-1.61	0.11	7.58	0.00	-4.39	**0.00**	Moderate
Hardware catalog 3	**0.14**	-1.95	**0.05**	3.70	0.00	-1.44	0.15	6.63	0.00	-6.14	**0.00**	Moderate
Industrial controls	0.00	-2.68	**0.01**	3.56	0.00	-2.74	**0.01**	7.74	0.00	-10.84	**0.00**	Alternatives
Personal care 1	**0.54**	-0.57	0.57	1.48	0.14	-0.58	0.56	3.42	0.00	-0.59	0.28	Good
Personal care 2	**0.41**	-1.01	0.31	-1.06	0.29	-1.23	0.22	4.21	0.00	-1.48	**0.09**	Moderate
Personal care 3	**0.86**	-0.04	0.97	4.30	0.00	0.87	0.38	4.54	0.00	-0.15	0.58	Good
Pharma 1	0.00	-12.96	**0.00**	-21.56	**0.00**	-11.81	**0.00**	18.40	0.00	-215.06	**0.00**	Alternatives
School biology labs	**0.10**	-1.25	0.21	4.26	0.00	5.80	0.00	5.17	0.00	-2.63	**0.02**	Moderate
School science labs	0.00	-3.36	**0.00**	5.64	0.00	-3.43	**0.00**	10.65	0.00	-16.68	**0.00**	Alternatives
Service parts 1, auto	0.00	-6.01	**0.00**	-6.41	**0.00**	-6.10	**0.00**	10.99	0.00	-51.99	**0.00**	Alternatives
Service parts 2, trucks	0.00	-3.29	**0.00**	0.29	0.77	-2.41	**0.02**	11.55	0.00	-16.26	**0.00**	Alternatives
Service parts 3, trucks	0.00	-7.72	**0.00**	25.50	0.00	21.86	0.00	22.33	0.00	-94.81	**0.00**	Alternatives
Service parts 4, trucks	**0.95**	-0.94	0.35	2.87	0.00	0.45	0.65	6.67	0.00	-1.62	**0.07**	Moderate
Soda 1	0.00	-2.73	**0.01**	1.21	0.23	1.09	0.28	5.68	0.00	-11.02	**0.00**	Alternatives

Table 10: Continued on next page

Table 10: Continued from previous page

Context	PL	Lognormal		Exp.		Str. Exp.		Poisson		PL+cut-off		Evidence
	p	LR	p	LR	p	LR	p	LR	p	LR	p	
Soda 2, small orders I	0.01	-2.74	**0.01**	-4.59	**0.00**	-2.92	**0.00**	5.88	0.00	-5.30	**0.00**	Alternatives
Soda 2, unit loads I	0.00	-4.22	**0.00**	-5.50	**0.00**	-4.50	**0.00**	14.87	0.00	-32.42	**0.00**	Alternatives
Supply 1, office	**0.18**	-1.46	0.15	4.90	0.00	-1.28	0.20	7.56	0.00	-3.78	**0.01**	Moderate
Supply 2, school	**0.93**	-0.29	0.77	2.16	0.03	1.26	0.21	4.67	0.00	-0.23	0.50	Good
Wine	0.00	-2.32	**0.02**	-2.03	**0.04**	-2.41	**0.02**	4.00	0.00	-7.54	**0.00**	Alternatives

Table 11: Tests of power law behavior and plausibility of alternative distributions in continuous data sets. For each data set, we give a p-value for the Power Law (PL) fit and the Log-likelihood Ratio (LR) for the alternative distributions considered: Lognormal, Exponential (Exp.), Stretched Exponential (Str. Exp.) and Power Law with Exponential Cut-off (PL+cut-off) distributions. We also include the p-values that indicate the significance of each of the likelihood ratio tests. Statistically significant p-values are in **bold**.

Context	PL	Lognormal		Exp.		Str. Exp.		PL+cut-off		Evidence
	p	LR	p	LR	p	LR	p	LR	p	
Grocery 1	**0.10**	-0.99	0.32	-0.44	0.66	-1.04	0.30	-1.63	**0.07**	Moderate
Grocery 2	0.03	0.53	0.60	2.48	0.01	0.68	0.50	0.00	1.00	None
Grocery 3, toxic	0.00	-17.69	**0.00**	-52.59	**0.00**	-16.01	**0.00**	-100.26	**0.00**	Alternatives
Pharma 2	**0.38**	-1.66	0.10	4.73	0.00	-1.70	0.09	-5.35	**0.00**	Moderate
Pharma 3	**0.15**	-1.12	0.26	-0.44	0.66	-0.25	0.81	-1.69	**0.07**	Moderate

6.6 Power Law generation mechanisms

As we noted earlier, power laws emerge from empirical distributions in economics and several other diverse areas. This widespread observed regularity has been explained in many ways. In this section, we review some of the proposed explanations for power laws and conjecture what might be applicable in the warehouse scenario.

In economics there have been, broadly described, two classes of explanations for the emergence of power laws. One is using models based on micro economic assumptions and the other based on stochastic models which seek to explain the observed distribution as a consequence of basic probabilistic assumptions. In this section, we primarily focus on the latter class of models that are called generation mechanisms to explain the empirical observations of the warehouse picks-per-SKU following power laws. Our approach is based on the belief that since power laws appear in so many varied situations that more general probabilistic mechanisms for power laws are more plausible than those that analyze scenario specific micro economic factors. Adopting probabilistic models does not imply that economic factors have no influence the behavior of these systems. In fact, the economic factors do bear influence on system behavior. However, these factors are so diverse in nature that their final combined effect may be modeled as a probability distribution and that their compound effect on the emergence of power laws may be viewed as essentially random components in a stochastic model. This is similar in the manner the Central Limit Theorem can be used to explain the widespread occurrence of the normal (Gaussian) distribution in nature.

In the next two sections, we consider two broad classes of stochastic models that have been used to explain power laws in other similar settings like city sizes, biological taxa among others.

6.7 *Preferential Attachment models*

The preferential attachment model is one of the earliest attempts at explaining the emergence of power laws. The original version of the model was proposed by G. Udny Yule using *pure birth* processes to explain the power law distribution of the sizes of biological taxa. Simon [37] refined Yule's approach and called it the *Preferential Attachment* model. Krugman's presentation [25] of Simon's preferential attachment model in the context of city sizes is as follows. In the model, the total population grows by constant increments, that Krugman [25] terms as *lumps.* At each time unit, a new lump is created and with probability π, forms a new city and with probability $1 - \pi$, attaches itself to an existing city with the probability that any particular city gets this lump being proportional to its population. With this setup, Simon assumes that the probability distribution of the population reaches steady state and derives a power law in the upper tail with a scaling parameter $\beta = 1 + \frac{1}{1-\pi}$. In the case of the urban population, from empirical studies we know that the scaling parameter is approximately 2 and to obtain that with this model, it is necessary that $\pi = 0$, meaning no new cities allowed to form. But this also leads to degeneracy in the model and hence the steady state assumption may not be valid [25]. On the other hand, if π is a small value close to 0, the model converges to a power law very slowly [25].

There have been other presentations of Simon's model with more mathematically precise arguments with the same result. The more general form is presented by Reed and Hughes [32] who extends this model to a birth and death process from a pure birth process. They model a homogeneous birth-and-death process with birth and death rates, λ and δ. If X_t is the population of a city at time t, the probabilities of a birth and death between time t and time $t + h$ is as follows:

$$P(X_{t+h} = n + 1 | X_t = n) = \lambda n h + o(h)$$

$$P(X_{t+h} = n - 1 | X_t = n) = \delta n h + o(h)$$

They also assume that new cities are formed in a similar process and hence the time since formation of each city is exponentially distributed at the rate ρ. They show that a power law emerges in the upper tail of the distribution of X as long as the birth rate is greater than the death rate, or, $\lambda > \delta$. The scaling parameter turns out to be $2 + \frac{\rho}{\lambda - \delta}$. This is a simple and promising generation mechanism for power laws that could potentially explain emergence of power laws in the picks-per-SKU per time period statistic in warehouses. The birth and death process accounts new SKUs that are introduced as well as old SKUs discontinued in the warehouse and the respective distributions appear reasonable ones to use in the aggregate. However the assumption that the birth rate is greater than death rate implies a warehouse whose overall number of picks is growing.

6.8 Multiplicative processes

In 1931, the French economist Gibrat proposed a simple model to explain the empirically observed size distribution of companies. He made the following assumptions: (i) the growth rate of a company is independent of its size, which is the *law of proportionate effect*, (ii) the successive growth rates are uncorrelated in time, and (iii) the companies do not interact. In mathematical form, if X_t is the size of a company at time t, Gibrat's model is expressed by the stochastic process:

$$X_{t+\Delta t} = X_t(1 + \epsilon_t). \tag{6.8.1}$$

where $X_{t+\Delta t}$ and X_t are, respectively, the size of the company at times $t + \Delta t$ and t, and ϵ_t is an uncorrelated random number with some bounded distribution. Hence $\log X_t$ follows a simple random walk [38], and after a sufficient time interval $t \gg 0$, the firm sizes X are lognormally distributed. These types of models are known as *multiplicative processes* and are used in many areas to explain the emergence of lognormal distributions.

Though Gibrat's model produces a lognormal distribution, with a minor change the model can serve as a mechanism to generate power laws. To generate a power law out of a multiplicative process, Mitzenmacher [26] cites two results of Champernowne and Kestner that state that a power law results if a lower bound is imposed on the random process $\{X_t\}$. Gabaix [17] demonstrates for a diffusion process (Brownian motion) based multiplicative model, imposing a lower bound, $x_{\min}$, results in a power law distribution and uses such a model to explain the Zipf law for city sizes.

Another approach is by Reed and Hughes [31] who impose a distribution on t, that can be interpreted as the lifetime or age of X_t. In a discrete multiplicative process, imposing a geometric distribution on t generates a power law distribution on X. Similarly, if X_t grows as a geometric Brownian motion, a type of multiplicative process, and assuming t, or the age of each X_i has an exponential distribution, then the resulting distribution on X is a power law.

This model and its variants has been used to explain power law distribution of computer file sizes [27] and also of the number of pages of World Wide Web sites [2]. The works of Gabaix [17] and Reed [30] use Brownian motion for the growth process which is a form of multiplicative process.

The multiplicative process also appears to be a plausible explanation of the power law observed in warehouses. However, this also requires that the total number of picks grows over time as it does in other settings where this model is used. It is entirely possible that year over year growth maybe an important factor that may decide whether a power law is observed in the data.

6.9 Conclusions

Power laws have been observed in many areas of science and warehouses appear to be no exception to this rule. In our analysis, we fit power laws to the picks-per-SKU statistic from 30 different warehouses and find statistical support in 16 of them. The power laws are also visually apparent in the upper tail region for these 16 data sets when we plot the cumulative complementary distribution function in log-log scales. We also notice warehouses in the Apparel, Hardware catalog, Personal care, Pharma, and Supplies categories show statistical support for the power law fit whereas the remaining do not, indicating a pattern in industry categories. However, the scaling parameter within each category appears to vary in a wide range and there appears to be no easy way of discerning industry from scaling parameter of the power law fit or the other way around.

Among the 16 data sets where the power law fit is plausible, we can rule out the alternative distributions only in 4 data sets. The remaining 12 show moderate support for the power law where an alternative distribution is considered more plausible. In all cases, the alternative distribution with greater statistical support is the power law with exponential cut-off, a generalization of the power law.

We also tested for alternative distributions and did indeed find the power law to be the best fit in 4 data sets, ruling out all alternative distributions. In 12 data sets where the power law is plausible, there is an alternative distribution that is viewed as more plausible than the power law. In all 12 of these cases the power law with an exponential cut-off as the most commonly plausible alternative distribution. However, this alternative is actually a generalization of the power law, meaning that statistical tests can identify this as a plausible alternative to the power law. These 12 data sets can be considered to have a *moderate* power law fit.

Among the remaining 14 data sets, 13 show evidence of at least one alternative distribution fitting the data set better than the power law and the power law with exponential cut-off is always one of them. The power law with exponential cut-off distribution has a thinner tail than the power law and implies that the large values of picks-per-SKU are less probable than the power law. Our hypothesis is that this could be because the warehouse has hit the limits of its market share and the demand for SKUs in the upper tail cannot grow further. This hypothesis corresponds to assumption required for power law generation mechanisms, which is perpetual growth. Based on this hypothesis, it is entirely possible that some warehouses in the growth phase with more market share to claim are candidates for a power law distribution. On the other hand, warehouses with saturated mature markets, may display distributions with thinner tails like the power law with exponential cut-off or even a lognormal distribution. With data from multiple years from warehouses at various stages of maturity, one can potentially verify this hypothesis.

REFERENCES

[1] Adamic, L. A., "Zipf, Power-laws, and Pareto - A ranking tutorial." www.parc.xerox.com/istl/groups/iea/papers/ranking/ranking.html.

[2] Adamic, L. A. and Huberman, B. A., "Zipf's law and the Internet," *Glottometrics*, vol. 3, no. 1, pp. 143–150, 2002.

[3] Ahuja, R. K., Orlin, J. B., Stein, C., and Tarjan, R. E., "Improved algorithms for bipartite network flow," *SIAM J. Comput.*, vol. 23, pp. 906–933, Oct. 1994.

[4] Anderson, C., *The Long Tail: How Endless Choice is Creating Unlimited Demand*. Random House, 2010.

[5] Bartholdi, J. J., "The knapsack problem," in *Building Intuition: Insights From Basic Operations Management Models and Principles* (Chhajed, D. and Lowe, T., eds.), International Series in Operations Research and Management Science, pp. 19–31, Springer Science, 2008.

[6] Bartholdi, J. J. and Hackman, S. T., "Simple ranking properties for the multi-mode forward-reserve problem." Georgia Institute of Technology.

[7] Bartholdi, J. J. and Hackman, S. T., "Warehouse and distribution science." `http://www.warehouse-science.com/book/wh-sci.pdf`.

[8] Bartholdi III, J. J. and Hackman, S. T., "Allocating space in a forward pick area of a distribution center for small parts," *IIE Transactions*, vol. 40, no. 11, pp. 1046–1053, 2008.

[9] Bozer, Y. A., *Optimizing throughput performance in designing order picking systems*. PhD thesis, 1985.

[10] Chekuri, C. and Khanna, S., "A PTAS for the multiple knapsack problem," in *Proceedings of the eleventh annual ACM-SIAM symposium on Discrete algorithms*, SODA '00, (Philadelphia, PA, USA), pp. 213–222, Society for Industrial and Applied Mathematics, 2000.

[11] Clauset, A., Shalizi, C., and Newman, M., "Power-law distributions in empirical data," *SIAM Review*, vol. 51, no. 4, pp. 661–703, 2009.

[12] Federgruen, A., Queyranne, M., and Zheng, Y. S., "Simple power-of-two policies are close to optimal in a general class of production/distribution networks with general joint setup costs," *Mathematics of Operations Research*, vol. 17, no. 4, pp. 951–963, 1992.

[13] Frazelle, E. H., *World-Class Warehousing.* Logistics Resources International, 1996.

[14] Frazelle, E. H., Hackman, S. T., Passy, U., and Platzman, L. K., "The forward-reserve problem," in *Optimization in industry* (Ciriani, T. A. and Leachman, R. C., eds.), vol. 2, pp. 43–61, New York, NY, USA: John Wiley & Sons, Inc., 1994.

[15] Fréville, A., "The multidimensional 0-1 knapsack problem: An overview," *European Journal of Operational Research*, vol. 155, no. 1, pp. 1–21, 2004.

[16] Frieze, A. and Clarke, M., "Approximation algorithms for the m-dimensional 0-1 knapsack problem: Worst case and probabilistic analysis," *European Journal of Operational Research*, vol. 15, pp. 100–109, 1984.

[17] Gabaix, X., "Zipf's laws for cities: An explanation," *The Quarterly Journal of Economics*, vol. 114, pp. 739–767, August 1999.

[18] Goldberg, A. V., "An efficient implementation of a scaling minimum-cost flow algorithm," *J. Algorithms*, vol. 22, pp. 1–29, Jan. 1997.

[19] Goldberg, A. V. and Tarjan, R. E., "Finding minimum-cost circulations by successive approximation," *Math. Oper. Res.*, vol. 15, pp. 430–466, July 1990.

[20] Gu, J., *The forward reserve warehouse sizing and dimensioning problem.* PhD thesis, 2005.

[21] Gu, J., Goetschalckx, M., and McGinnis, L. F., "Solving the forward-reserve allocation problem in warehouse order picking systems," *Journal of the Operational Research Society*, vol. 61, no. 6, pp. 1013–1021, 2010.

[22] Hackman, S. and Platzman, L., "Near-optimal solution of generalized resource allocation problems with large capacities," *Operations Research*, vol. 38, pp. 902–910, 1990.

[23] Hackman, S. T. and Rosenblatt, M. J., "Allocating items to an automated storage and retrieval system," *IIE Transactions*, vol. 22, pp. 7–14, March 1990.

[24] Kellerer, H., Pferschy, U., and Pisinger, D., *Knapsack Problems.* Springer, 2004.

[25] Krugman, P., "Confronting the mystery of urban hierarchy," *Journal of the Japanese and International Economies*, vol. 10, pp. 399–418, 1996.

[26] Mitzenmacher, M., "A brief history of generative models for power law and lognormal distributions," *Internet Mathematics*, vol. 1, no. 2, pp. 226–251, 2004.

[27] Mitzenmacher, M., "Dynamic models for file sizes and double Pareto distributions," *Internet Mathematics*, vol. 1, no. 3, pp. 305–333, 2004.

[28] Newman, M. E., "Power laws, Pareto distributions and Zipf's law," *Contemporary physics*, vol. 46, no. 5, pp. 323–351, 2005.

[29] Orlin, J. B., "A faster strongly polynomial minimum cost flow algorithm," *Operations Research*, vol. 41, pp. 338–350, March-April 1993.

[30] Reed, W. J., "On the rank-size distribution for human settlements," *Journal of Regional Science*, vol. 41, pp. 1–17, 2002.

[31] Reed, W. J. and Hughes, B. D., "From gene families and genera to incomes and internet file sizes: Why power laws are so common in nature," *Physical Review E*, vol. 66, no. 6, 2002.

[32] Reed, W. J. and Hughes, B. D., "Power-law distributions from exponential processes: An explanation for the occurrence of long-tailed distributions in biology and elsewhere," *Scientiae Mathematicae*, vol. 8, 2003.

[33] Roundy, R., "98%-effective integer-ratio lot-sizing for one-warehouse multi-retailer systems," *Management Science*, vol. 31, no. 11, pp. pp. 1416–1430, 1985.

[34] Roundy, R., "A 98%-effective lot-sizing rule for a multi-product, multi-stage production/inventory system," *Mathematics of Operations Research*, vol. 11, no. 4, pp. pp. 699–727, 1986.

[35] Roundy, R., "Rounding off to powers of two in continuous relaxations of capacitated lot sizing problems," *Management Science*, vol. 35, no. 12, pp. pp. 1433–1442, 1989.

[36] Sen, P. K. and Singer, J. M., *Large Sample Methods in Statistics*. Chapman and Hall, 1993.

[37] Simon, H., "On a class of skew distribution functions," *Biometrika*, vol. XLII, pp. 425–440, 1955.

[38] Sutton, J., "Gibrat's legacy," *Journal of Economic Literature*, vol. 35, pp. 40–59, March 1997.

[39] van den Berg, J. P., Sharp, G. P., Gademann, A., and Pochet, Y., "Forward-reserve allocation in a warehouse with unit-load replenishments," *European Journal of Operational Research*, vol. 111, no. 1, pp. 98–113, 1998.

[40] Walter, R., Boysen, N., and Scholl, A., "The discrete forward-reserve problem - Allocating space, selecting products, and area sizing in forward order picking," *European Journal of Operational Research*, vol. 229, no. 3, pp. 585–594, 2013.